CUK-H06251-25-P16654

THE
BRITISH JOURNAL
OF
PSYCHOLOGY

25

CAMBRIDGE UNIVERSITY PRESS
LONDON: BENTLEY HOUSE, N.W. 1

CHICAGO: THE UNIVERSITY OF CHICAGO PRESS
(Agents for the United States)
BOMBAY, CALCUTTA, MADRAS: MACMILLAN & CO., LTD.
TOKYO: MARUZEN COMPANY, LTD.

All rights reserved

THE
BRITISH JOURNAL
OF
PSYCHOLOGY

GENERAL SECTION

25

EDITED BY
F. C. BARTLETT
WITH THE ASSISTANCE OF
C. BURT, J. DREVER, AND S. S. ISAACS

AIDED IN THE SELECTION OF PAPERS BY

F. AVELING
E. A. BOTT
W. BROWN
B. EDGELL
G. DAWES HICKS
J. T. MacCURDY
C. S. MYERS
T. H. PEAR
C. S. SHERRINGTON
C. SPEARMAN
G. UDNY YULE

VOLUME XXIX, 1938–39

CAMBRIDGE
AT THE UNIVERSITY PRESS
1939

P16654

PRINTED IN GREAT BRITAIN

CONTENTS OF VOL. XXIX

PART 1. JULY, 1938

PAGE

The Method of testing Literary Appreciation. By C. FOX . 1

The Acquisition of Verbal Repetition Habits. By R. C. OLDFIELD and O. L. ZANGWILL 12

The Measurement of the Threshold for Flicker and its Value as a Perseveration Test. By S. BIESHEUVEL 27

The Effect of Knowledge of Results on Learning and Performance. I. A Co-ordinated Movement of the Two Hands. By J. L. ELWELL and G. C. GRINDLEY 39

Speed of Work in Intelligence Tests. By P. SLATER . . . 55

Note on Professor Godfrey H. Thomson's Article "The Influence of Univariate Selection on Factorial Analysis of Ability". By W. LEDERMANN 69

Intelligence, Concrete and Abstract: Note. By W. P. ALEXANDER 74

Critical Notice: *Psychology down the Ages*. By C. Spearman. By O. L. ZANGWILL 75

Publications Recently Received 79

British Psychological Society. Names of Officers and Members i

British Psychological Society. Rules xxiii

British Psychological Society. Library Rules and List of Periodicals xxix

PART 2. OCTOBER, 1938

PAGE
The Effect of Veiling Glare on Apparent Size Relations. By T. MURRAY MARTIN and R. W. PICKFORD 91
Experiments on the Constancy of Shape. By W. E. MOORE . 104
A Note on an Improved Design of Voice Key. By D. F. VINCENT 117
Characteristic Motivation in the Activities of School-Girls. By M. D. VERNON 121
Vocational Preferences of School-leavers in a Scottish Industrial Area. By H. PALLISTER 144
The Ultimate Basis of Aesthetic Experience. By W. R. D. FAIRBAIRN 167
Proposed Explanation of Individual Differences of Ability by "Sampling". By C. SPEARMAN 182
Publications Recently Received 192

PART 3. JANUARY, 1939

Obituary Notice. Professor Wilhelm Stern, 1871–1938 . . 205
The Co-operation of Social Groups. Part I. By D. M. CARMICHAEL 206
Characteristic Motivation in the Activities of School-Girls. By M. D. VERNON 232
A Preliminary Note on a Vocational Investigation. By ALICE HEIM 246
The Effect of Adaptation upon Visual Acuity. By K. J. W. CRAIK 252
Reliability of Mental Tests. By ROBERT W. B. JACKSON . 267
The Influence of Multivariate Selection on the Factorial Analysis of Ability. By GODFREY H. THOMSON and WALTER LEDERMANN 288
Editor's Note 306
Publications Recently Received 307
British Psychological Society—Revenue Account and Balance Sheet 315

PART 4. APRIL, 1939

PAGE
Obituary Notice. Professor Samuel Alexander, 1859–1938 . 317
„ „ Professor William McDougall, 1871–1938 . 321
The Co-operation of Social Groups. Part II. By D. M. CARMICHAEL 329
Variability of Judgement and 'Steadiness' of Character. By JEAN D. CUMMINGS 345
Some Relations between Reproducing and Recognizing Prose Material. By O. L. ZANGWILL 370
The Effects of Errors of Measurement on Correlation Coefficients. By ROBERT H. THOULESS 383
The General Factor in Correlations between Persons. By M. DAVIES 404
Temperament and the Threshold for Flicker. By LL. WYNN JONES 422
A Further Note on the Degree of Variation in the Score of a Motor Test of Perseveration. By JANE DARROCH . . 427
Publications Recently Received 429
Fifth International Congress for the Unity of Science. Preliminary Announcement 437
Proceedings of the British Psychological Society . . . 438

LIST OF AUTHORS

PAGE
Alexander, W. P. Intelligence, Concrete and Abstract: Note 74
Bieshevel, S. The Measurement of the Threshold for Flicker and its Value as a Perseveration Test 27
British Psychological Society. Library Rules and List of Periodicals xxix
British Psychological Society. Names of Officers and Members i
British Psychological Society—Revenue Account and Balance Sheet 315
British Psychological Society. Rules xxiii
Carmichael, D. M. The Co-operation of Social Groups. Part I 206
Carmichael, D. M. The Co-operation of Social Groups. Part II 329
Craik, K. J. W. The Effect of Adaptation upon Visual Acuity . 252
Cummings, Jean D. Variability of Judgement and 'Steadiness' of Character 345
Darroch, Jane. A Further Note on the Degree of Variation in the Score of a Motor Test of Perseveration . . . 427
Davies, M. The General Factor in Correlations between Persons 404
Editor's Note 306
Elwell, J. L. and Grindley, G. C. The Effect of Knowledge of Results on Learning and Performance. I. A Co-ordinated Movement of the Two Hands 39
Fairbairn, W. R. D. The Ultimate Basis of Aesthetic Experience 167
Fifth International Congress for the Unity of Science. Preliminary Announcement 437
Fox, C. The Method of testing Literary Appreciation . . 1
Heim, Alice. A Preliminary Note on a Vocational Investigation 246
Jackson, Robert W. B. Reliability of Mental Tests . . 267

PAGE

JONES, LL. WYNN. Temperament and the Threshold for Flicker 422

LEDERMANN, W. Note on Professor Godfrey H. Thomson's Article "The Influence of Univariate Selection on Factorial Analysis of Ability" 69

MARTIN, T. MURRAY and PICKFORD, R. W. The Effect of Veiling Glare on Apparent Size Relations 91

MOORE, W. E. Experiments on the Constancy of Shape . 104

Obituary Notice. Professor Samuel Alexander, 1859–1938 . 317

„ „ Professor William McDougall, 1871–1938 . 321

„ „ Professor Wilhelm Stern, 1871–1938 . . 205

OLDFIELD, R. C. and ZANGWILL, O. L. The Acquisition of Verbal Repetition Habits 12

PALLISTER, H. Vocational Preferences of School-leavers in a Scottish Industrial Area 144

Proceedings of the British Psychological Society . . . 438

Publications Recently Received . . . 79, 192, 307, 429

SLATER, P. Speed of Work in Intelligence Tests . . . 55

SPEARMAN, C. Proposed Explanation of Individual Differences of Ability by "Sampling" 182

THOMSON, GODFREY H. and LEDERMANN, WALTER. The Influence of Multivariate Selection on the Factorial Analysis of Ability 288

THOULESS, ROBERT H. The Effects of Errors of Measurement on Correlation Coefficients 383

VERNON, M. D. Characteristic Motivation in the Activities of School-Girls 121

VERNON, M. D. Characteristic Motivation in the Activities of School-Girls 232

VINCENT, D. F. A Note on an Improved Design of Voice Key 117

ZANGWILL, O. L. Critical Notice: *Psychology down the Ages.* By C. Spearman 75

ZANGWILL, O. L. Some Relations between Reproducing and Recognizing Prose Material 370

Incomplete

105-
.003

VOLUME XXIX   OCTOBER, 1938 - 39 April)   PART 2

THE EFFECT OF VEILING GLARE ON APPARENT SIZE RELATIONS

By T. MURRAY MARTIN and R. W. PICKFORD

From the Psychological Laboratory, Glasgow University

I. *Introduction* (pp. 91–92).
II. *Method* (pp. 92–93).
III. *Statistical results* (pp. 93–97).
IV. *Descriptive results* (pp. 97–99).
V. *Discussion and conclusions* (pp. 99–102).
VI. *Summary* (pp. 102–103).
References (p. 103).

I. INTRODUCTION

If two precisely similar disks are placed upright before a subject, side by side, at the same distance from him, they look the same size. If one of them is now moved away from him in the line of vision, it usually begins to look smaller. The rate of decrease in its apparent size, however, is less than would be expected if apparent size were to follow the law of inverse squares. Thouless has used the expression "phenomenal regression to the real size", for the tendency of the subject to see the two disks more nearly the same size than the inverse square law would predict (3). Koffka gives an important account of the different aspects of this problem (2), and Miss Vernon mentions most of the relevant literature and gives an interesting discussion of it (4).

Under fog conditions the 'articulation' of the field of vision is reduced in varying degrees, and it would then be expected that the change in apparent size would more nearly follow the inverse square law than in normal vision. An approximation to fog conditions may be made by placing before the subject's eye a piece of plate glass at an angle of 45° to the line of vision, by which diffused light is reflected into the eye. The intensity of the diffused light is then inversely proportional to the square of the distance of the light source from the diffusing screen. This diffused light makes a veiling glare, much like the glare produced by a fog, when the subject looks into it.

In the present experiment the questions were asked, whether the effects of articulation of the visual field could be cut out progressively by first reducing binocular to monocular vision and then increasing veiling glare produced in the way described above, and what effects these steps would have on the relative apparent sizes of two disks at different distances. The experimenters were forced to ask also whether the instructions, to make the two disks look the same size, had the same meaning for all subjects.

II. METHOD

An apparatus was set up (Fig. 1) such that on two small white screens at different distances from the eye circular 'disks of light' could be projected, that on the nearer screen being of variable diameter, and that on the farther screen being of fixed size. The variable disk of light was produced by a diaphragm, placed in the lantern, the size of which could be varied by the subject's manipulating a loop of string. The screens were so mounted that the disks of light appeared at different distances along a white 'roadway', which gave an exaggerated perspective effect. This roadway was bordered by black cloth and the farther disk was seen against a black back-cloth. The roadway was illuminated by a 40 W. lamp mounted about 6 ft. above it. The roadway and the disks could be viewed through short black eye-tubes in one of which could be produced glare of varying intensity. The experimenter was able to read the size of the variable disk to half a millimetre, for in the beam of light giving the

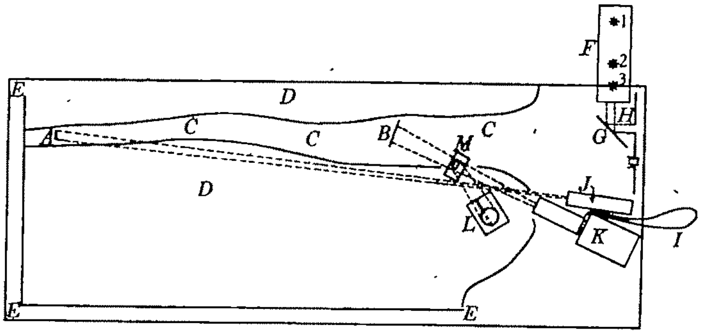


Fig. 1. *A*, standard disk; *B*, variable disk; *CC*, roadway; *D*, black cloth; *EE*, black screen; *F*, lamp-holder giving three positions (1, 2, 3) of lamp for glare; *G*, plate glass partial reflector; *H*, screen with two eye-holes, right temporarily blocked; *I*, loop of cord for subject's control of lantern diaphragm; *J*, spot-light for constant disk; *K*, lantern with variable diaphragm immediately in front of condensing lens; *L*, frosted glass plate with scale; *M*, plate glass partial reflector.

disk was placed a small glass plate which reflected an image on to a piece of frosted glass carrying a scale. When the image on the frosted glass plate was in focus, its size was necessarily the same as that of the variable disk, also in focus, seen by the subject. Three intensities of veiling glare were used, G_1, G_2, G_3, such that $G_3 - G_2$ was approximately equal to $G_2 - G_1$, and G_3 was of such a magnitude that only the disks of light could be seen through it.

The subject, in accordance with instructions to be mentioned below, made adjustments of equality under the three veiling glare conditions. In addition he made similar adjustments with no glare both monocularly (*N.G.m.*) and binocularly (*N.G.b.*). The amount of adjustment necessary, and its direction, were varied by the experimenter from reading to reading in the usual way. The series of conditions for one complete set of readings was G_1, G_2, G_3, *N.G.m.*, *N.G.b.*, in that order, and ten settings were taken for each of these conditions. One or more sets of readings were made by each subject in a day's sitting, as convenient, after a preliminary practice set.

The instructions given to the subject in the preliminary practice set were: "Adjust the size of the nearer disk until in your opinion it looks the same size as the farther disk." After this set of observations had been made by each subject he was asked to describe the method used to equate the apparent sizes of the disks under the five different conditions. It was found that the subjects were using different methods even though the same instructions were given to them all. Of the seven subjects taking part in the experiment, Subjects III and V assumed that it was necessary to think of the disks as if in the same plane, while the other subjects assumed that the disks were to be compared as if in different planes. The experimenters had intended to see how far the progressive reduction of 'articulation' of the field of vision would affect the actual sizes of the disks when appearing the same size at different distances. Now it became necessary also to study the different responses of the subjects to the same instructions. Five subjects were able to do the experiment with modified instructions, trying now to think of the disks in whichever way was the reverse of their previous method. The other two subjects were not accessible long enough for this to be done.

III. STATISTICAL RESULTS

Table I shows the numerical results for the five subjects who did both forms of the experiment. Subjects I, II and IV assumed that to make the disks look the same size meant seeing them as if at different distances.

Table I

Condition	Exp. I 'Same planes'					Exp. II 'Different planes'					Diff. of Exps. I and II	
	No.	Mean	S.E.	Diff.	S.E.*	No.	Mean	S.E.	Diff.	S.E.*	Diff.	S.E.*
SUBJECT I	(Done first)					(Done second)						
N.G.b.	30	18·3	0·19			50	14·9	0·11			3·4	0·22
				2·7	0·31				2·0	0·17		
N.G.m.	30	15·6	0·24			50	12·9	0·13			2·7	0·28
				2·8	0·29				1·7	0·17		
G_1	30	12·8	0·16			50	11·2	0·11			1·6	0·17
				2·1	0·21				1·5	0·16		
G_2	30	10·7	0·14			50	9·7	0·12			1·0	0·15
				2·2	0·16				1·5	0·14		
G_3	30	8·5	0·09			50	8·2	0·08			0·3	0·12
SUBJECT II	(Done first)					(Done second)						
N.G.b.	30	20·1	0·14			50	17·4	0·10			2·7	0·17
				1·5	0·24				1·6	0·16		
N.G.m.	30	18·6	0·19			50	15·8	0·12			2·8	0·22
				4·6	0·25				2·5	0·16		
G_1	30	14·0	0·16			50	13·3	0·10			0·7	0·19
				1·9	0·25				1·0	0·17		
G_2	30	12·1	0·19			50	12·3	0·13			−0·2	0·23
				2·5	0·26				4·0	0·23		
G_3	30	9·6	0·18			50	8·3	0·19			1·3	0·26
SUBJECT III	(Done second)					(Done first)						
N.G.b.	30	12·3	0·10			50	11·0	0·07			1·3	0·12
				1·4	0·12				1·1	0·09		
N.G.m.	30	10·9	0·02			50	9·9	0·06			1·0	0·06
				0·2	0·12				0·2	0·12		
G_1	30	10·7	0·10			50	9·7	0·10			1·0	0·14
				0·4	0·15				0·6	0·12		
G_2	30	10·3	0·11			50	9·1	0·06			1·2	0·12
				1·6	0·13				0·8	0·10		
G_3	30	8·7	0·07			50	8·3	0·08			0·4	0·12
SUBJECT IV	(Done first)					(Done second)						
N.G.b.	50	13·8	0·07			50	11·3	0·10			2·5	0·12
				1·3	0·11				0·7	0·13		
N.G.m.	50	12·5	0·09			50	10·6	0·09			1·9	0·13
				0·6	0·13				0·3	0·10		
G_1	50	11·9	0·10			50	10·3	0·05			1·6	0·11
				0·1	0·16				0·5	0·09		
G_2	50	11·8	0·12			50	9·8	0·08			2·0	0·13
				2·0	0·13				0·8	0·13		
G_3	50	9·8	0·06			50	9·0	0·10			0·8	0·12
SUBJECT V	(Done second)					(Done first)						
N.G.b.	20	15·8	0·11			50	15·0	0·14			0·8	0·24
				2·1	0·15				1·8	0·21		
N.G.m.	20	13·7	0·10			50	13·2	0·16			0·5	0·26
				0·1	0·18				0·2	0·23		
G_1	20	13·6	0·15			50	13·0	0·16			0·6	0·27
				1·7	0·18				0·9	0·21		
G_2	20	11·9	0·09			50	12·1	0·14			−0·2	0·24
				2·1	0·17				1·8	0·18		
G_3	20	9·8	0·15			50	10·3	0·12			−0·5	0·20

* The S.E.'s for Subject V were estimated by Fisher's method for differences between small samples(1).

They did Exp. I first and II second. Subjects III and V did Exp. II first, because they assumed that the disks would have to be thought of in the same plane, and Exp. I second. The various mean readings for the numbers of observations shown are given in millimetres, together with their standard errors. All these means are statistically significant. It will be seen that the highest values of the variable for these five subjects are with binocular vision and no glare. The standard was 18 mm. in diameter and at a distance of 3 m. from the subject's eye. The variable was at 1·34 m. The highest mean value of the variable, if we leave out Subject VII, who will be considered later, was 20·1 mm., given by Subject II (*N.G.b.*; Exp. I). The lowest mean value was 8·2 mm., given by Subject I (G_3; Exp. II) (see Fig. 2), and is very near the size expected according to the law of inverse squares (8·04 mm.). The differences between the mean values of the variable under the five conditions of vision, with and without veiling glare, are also shown, with their standard errors. Thirty-five out of the forty differences may be taken as significant. For all five subjects the change from *N.G.b.*, which is normal vision, to G_3, in which distance cues are completely cut out, means a reduction in the size of the variable which is chosen as the same as the standard, more and more approaching the size expected on the inverse square law.

This table also shows the differences in the mean values of the variables obtained under the same conditions of vision but by the different methods of Exps. I and II. These differences are all significant, excepting those of Subjects II and V under the condition G_2, both of which are reversals, and of Subject V under *N.G.m.* Of the other results of Subject V under glare, one is doubtful and the other is a reversal. Where the results are statistically significant, it is concluded that the change in the subject's method of working had a measurable effect. It is of interest that all differences under normal binocular vision are significant, and that the magnitudes of the differences decrease, on the whole, from *N.G.b.* towards G_3. This would be expected, because the less distance cues operate the easier will it be to think of constant and variable disks as if in the same plane.

Table II shows the results for Subjects VI and VII, who were not able to do the experiment with modified instructions. Although Subject VI did only ten readings under each condition, her results support the general conclusions, but show one reversal. Subject VII, however, gave very remarkable results. For him there is a significant difference between the mean readings for *N.G.b.* and *N.G.m.*, and in the usual direction, but the differences between the mean readings under the three degrees of veiling glare are not significant. The difference between *N.G.m.* and G_1 are significant, but reversed: he has given a greater value to the variable—already very large—when veiling glare was present, and the same value whatever the glare. This suggests that for him the distant disk looked

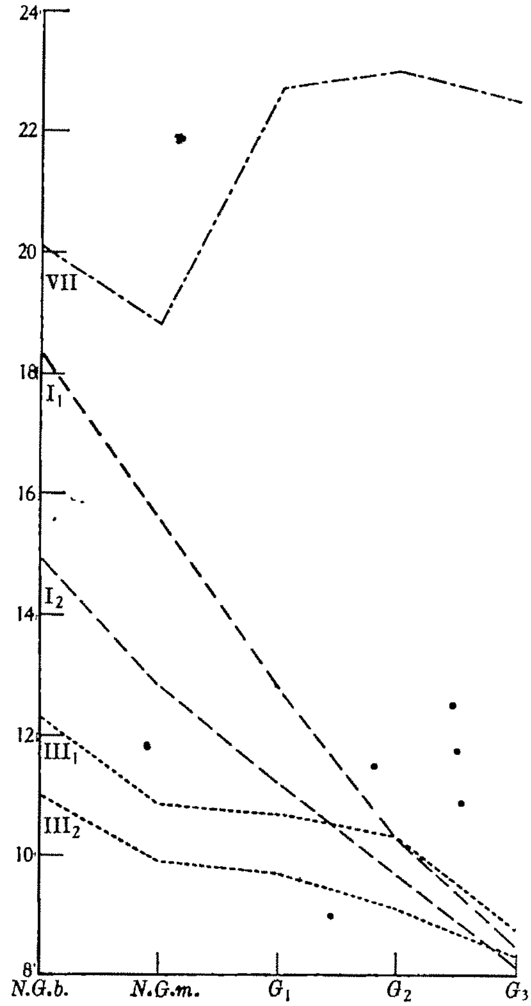


Fig. 2. Results for Subject I, Exps. I and II; Subject III, Exps. I and II; and for Subject VII. The distances between *N.G.b.* and *N.G.m.* are not mathematically meaningful, but those between G_1, G_2, and G_3 are approximately proportional to the corresponding changes in intensity of glare.

slightly larger under glare and not relatively smaller as for the other subjects.

Table II. *Subjects VI and VII*

Condition	Subject VI					Subject VII				
	No.	Mean	S.E.	Diff.	S.E.*	No.	Mean	S.E.	Diff.	S.E.*
N.G.b.	10	15·0	0·18			30	21·0	0·30		
				3·0	0·19				2 2	0·47
N.G.m.	10	12 0	0·05			30	18·8	0·36		
				−0·5	0·13				−3·9	0·50
G_1	10	12·5	0·12			30	22·7	0·34		
				1·5	0·23				−0·3	0·51
G_2	10	11·0	0·20			30	23·0	0·37		
				1·5	0·47				0·5	0·65
G_3	10	9·5	0·43			30	22·5	0·53		

* Estimated S.E.'s by Fisher's method for differences between small samples(1).

IV. DESCRIPTIVE RESULTS

It is important to consider the methods of the subjects more carefully, because of the possible assumption that the request to make the two disks look the same size can have only one meaning. Subjects I, II and IV realized that this experiment was on apparent size relations at different distances, and took no steps which might weaken the influence of apparent distance. Of these Subjects I and IV were fully aware of the psychological implications of the experiment and Subject II was moderately acquainted with them.

Subject I could easily distinguish two methods of equating the disks which gave, in preliminary experiments, different results: (1) Equating by what he called the "Total response method". "This can only be described as responding to the 'size-in-its-setting' of the further disk, and then transferring this to the nearer disk, and trying to fit the second situation 'as a whole' to the first." Those who have taken part in experiments on size or brightness 'constancy' will realize what is meant by this statement. (2) "Method of reduction to the same plane." Here the result was different according to whether the two disks were imagined to be in the plane of the nearer or of the farther disk. No doubt two results could be got by the first method if 'one-way' comparison were used instead of rapid alternation from one disk to the other. If this is so, then the result got by this method may be a mean of the two modes of one-way comparison. "In my experiments", he said, "I used method 1 above, and the first part of method 2, reducing to the plane of the nearer disk." Subject I has shown a high phenomenal regression in previous experiments.

For Subject IV equalizing the disks was a considerable problem in binocular vision without glare. He appears to have had the least degree of phenomenal regression of the three. For him the expression, "Look the same size", was almost without meaning in *N.G.b.* "The disks did not look the same size at all." Each was a "disk-of-a-certain-size-and-distance". "It was necessary to make the near disk such a size that it would look the same size as the distant disk if it were at the same distance." Expressed verbally, this process, which is rapid and immediate in character, seems absurdly round-about. With the reduction of distance effects, as the glare increased, the difficulty decreased, and the process became direct comparison. At the same time, however, the disks seemed less and less to "have distance" at all, and in G_3 they were simply like moons in a misty sky, just as far or as near as he liked to imagine. This subject also had difficulty with Exp. II, because he was much affected by the combination, "size-at-a-certain-distance", and, in order to think of the disks as if at the same distance, he had to imagine that the 'roadway' curled back upon itself and became a sort of overhanging arm which held the constant disk at the same distance as the variable. When this visual interpretation is not maintained, the constant disk "suddenly changes and becomes larger".

Subjects III and V assumed that to make the disks look the same size, as instructed, they would have to think of them as if in the same plane. Subject III was familiar with the problem of phenomenal regression, but simply took it for granted that her interpretation of the instructions was correct. In Exp. II (which she did first) "the variable was adjusted by looking rapidly from variable to constant and not allowing constant to fit into its proper surroundings and distance, i.e. by attempting to focus on constant as though it were at the same distance as variable. If one did allow constant to sink into its proper environment and distance then it appeared to expand and stand out in its distant position", and this gave the adjustment for Exp. I, "where attention was concentrated on constant which was focussed at its proper distance in its surroundings and variable was adjusted then to this size—occasionally constant appeared to recede into the remote distance and then it seemed to become much smaller". In Exp. I (done second) she explained the following points: "Sometimes far one recedes into distance and looks smaller—keep them about same distance"; "Vaguer to do this than other experiment"; "In other, aiming at something definite though more difficult"; "Gaze at standard and get idea of it with its distance; then try and make this one same size—look back to check up"; "Not merely comparison of two as in other experiment"; "Combination of size it is and distance it is." Subject V was not familiar with the psychological meaning of the experiments. He pointed out, however, in Exp. II, which he did first, that it is "Funny without glare—move eyes backwards and forwards; with glare both look in the same plane".

Subject VII, who chose a slightly larger variable under the veiling glare than in normal vision, reported that it was "Easier under glare". "Glare on nearer one gives wrong impression of size—making it a lot larger than it really is. At times I was judging them to be the same size but I was aware I was wrong with the near one—the glare was making them the same size. I had impression that glare increased size of near disk from reality. This would have led me to make it smaller than it ought to be. I watched against that and tried to judge them just as they appeared." For him there seem to be two possibilities: he might make them the same size according to what he calls 'reality', and then they would be as near as possible the same diameter, appearances altogether apart; or he might 'judge' them to appear the same size whatever their 'real' diameters might be. Under glare he continued to act upon appearances, as instructed, though he was aware of being "wrong with the near one".

V. DISCUSSION AND CONCLUSIONS

(1) It is not desirable to draw sweeping conclusions from a small experiment, but several points of great interest arise here. The numerical results show that under veiling glare there is approximation to the inverse square law for apparent size relations of objects seen at different distances. In this experiment the brightness relations under different conditions of glare, which might be affected in fog vision, were kept constant; the disks were always of the same apparent brightness. The articulation of the visual field is reduced by glare and the problem of making the two disks 'look' the same size is more like direct comparison of objects at the same distance. It would be reasonable to expect that in motoring under fog conditions phenomenal regression is reduced. Anyone with a strong phenomenal regression, who had no difficulty in seeing correctly that his car would pass through a given gap many yards ahead, might make serious mistakes in a fog, even when he felt fairly confident. This requires to be tested out under everyday conditions. It should be remembered that persons are accustomed to see for everyday purposes by the modes of vision that they actually use, whatever these may be, and that a man who had low phenomenal regression might be as well able to see size

relations under normal vision as a man with high regression, but would be far less disturbed in a fog. In view of this it is interesting that Subject IV, who had low phenomenal regression, reported that he had been surprised sometimes at shooting a rabbit running in the misty twilight, when he thought the bad conditions of vision would have made aiming almost impossible. Probably a person with very high phenomenal regression would have been much more disturbed by the reduced distance effects and quite unable to hit the mark.

Neither of these applications of the results of the experiment would hold for any person who saw the distant disk larger under veiling glare, and of relatively constant size whatever the glare. Subject VII appeared to compensate for the effects of glare, and would, presumably, see objects at a given distance slightly larger under fog conditions, however bad the fog. It is interesting to recall that the order of experimental procedure placed normal binocular vision last, and it is difficult to see how his compensation could have been made on conscious knowledge of the 'real' size of the standard. According to his own account, however, he could have worked on the basis of 'real' rather than 'apparent' sizes, and then would have chosen a smaller variable. It is very interesting to notice that the other subjects did not make such a sharp distinction between 'appearance' and 'reality', because they chose, on the ground of 'appearance', the size this subject would possibly have chosen on the ground of 'reality'. It is also very unfortunate that the experiment could not be continued at the time with him, to see what size he would indeed choose on the ground of 'reality'.

In general the results suggest that considerable individual differences must be expected in a large population. These differences might be important in testing motor drivers, air pilots and other people who had to take account of size relations in fog.

It should be pointed out that in fog people commonly observe that objects which can be seen look unexpectedly large, and this is due to the increase in their apparent distance. The numerical results for Subjects I–VI do not accord with this observation, but those of Subject VII accord with it completely. The whole matter obviously needs fuller investigation. Artists habitually use clearness of outline and presence of detail to make objects which they draw look near, and if they wish to make an object in the foreground specially unobtrusive, they take care that it does not have a clear outline or much detail of structure. It is probable that, in normal vision, sharpness of detail and of outline are important factors in determining apparent distance, but, in this experiment, the disks had no detail of visual structure at any time, and the veiling glare did not make the outline of the distant disk more blurred than of the near disk, as it would have been in a fog. Another interesting point is that very short preliminary observations on another subject have indicated no change whatever in phenomenal regression, when illumination was decreased until disks at different distances and not illuminated by special spotlights, were almost invisible. When this is done the outline of the distant disk does become much more blurred than that of the near disk.

(2) Turning to the question of experimental instructions, it might be surprising to some people that two subjects thought they ought to try to reduce the articulation of the visual field for themselves. It is, however, not so surprising as it seems, because to these subjects the assumption was reasonable that the two disks would be more satisfactorily made to 'look' the same size if thought of in the same plane. Indeed, they both found direct comparison difficult when the disks did not seem in the same plane, and adopted what seemed to them the best solution of the problem, to make the two disks look the same size.

The three subjects who did nothing to weaken the influence of distance cues, as it happens, were interested in the meaning of the experiment. They all spontaneously adopted what might be called the 'proper' attitude. The question of what attitudes are adopted in solving the problem of equalizing apparent sizes of similar disks seen at different distances is of great psychological interest. It is sometimes assumed that any instructions which tend to set up an 'analytic' attitude are psychologically illegitimate, and that the 'natural' attitude is satisfactorily set up by asking the subject simply to make the disks 'look' the same size. The present experiment shows that such instructions may set up the attitude thought proper to the experiment by the subject in question, and what one person may think proper another may not. More explicit instructions are required, and, indeed, they should probably be adapted to each individual subject. In other words, quantitative results, even though statistically significant, may be psychologically meaningless unless accompanied by adequate statements or remarks by the subjects, or unless the instructions have been carefully adapted to each subject and we know what they were. It is, indeed, most unlikely that psychology can become an exclusively quantitative science. This is important, because in psychology we always have to take account of the living—and often of the conscious—organism, and to consider all aspects of its response under given conditions. It is therefore well to think of psychology as a biological and not as a physical science.

A final word concerns the meaning of the words 'look the same These words have no absolute meaning, even in the setting of a standardized experiment. Their meaning is relative to the subject's idea of the situation in which he is placed. Two very different sizes of variable disk may both 'look' the same as the standard, if different interpretations are made of the situation. Under normal conditions, moreover, it may be doubtful, sometimes, whether two disks at different distances can be said strictly to 'look' the same at all. The subject may have to make an estimate of the size of one disk which would 'look' the same as the size of the other if it were at the same distance, without ever truly seeing them the same.

VI. SUMMARY

1. An experiment was done in which the sizes of two similar disks, one 18 mm. diameter, at 300 cm. distance, and the other variable in diameter, at 134 cm. distance, were equalized by the subject, in normal binocular vision, in normal monocular vision, and in monocular vision in three degrees of veiling glare. The strongest glare cut out all visual distance cues. It was found that, as distance cues are gradually cut out and glare increases, the apparent size relations of the two disks more and more nearly approach and finally almost reach what would be expected on the law of inverse squares. This reduction of phenomenal regression under veiling glare might be of interest in relation to motor driving and flying in fog. The experiment indicates that large individual differences in this reduction are to be found.

2. Three subjects assumed that the instructions, to make the two disks look the same size, meant thinking of the disks as if at different distances, but two subjects assumed that they would have to think of the disks as if in the same plane. Two other subjects were not available long enough to use more than one method, but the five mentioned repeated the experiment reversing the methods used at first. There was a large proportion of significant differences in the mean readings by the two methods. This suggests that it is not sufficient to assume that simple instructions are always interpreted in the way the experimenter expects. The subject responds in the way fitting for the conditions and instructions as he understands them. It is also possible that some subjects may doubt that two disks at different distances can 'look' strictly the same size in normal vision at all.

Our thanks are due to the subjects of the experiment, to Dr R. H. Thouless for his advice and criticism, and to Mr J. F. Simpson.

REFERENCES

(1) FISHER, R. A. (1936). *Statistical Methods for Research Workers*, 6th ed. Edinburgh; Oliver & Boyd.

(2) KOFFKA, K. (1935). *Gestalt Psychology*. London: Kegan Paul, etc.

(3) THOULESS, R. H. (1931). "Phenomenal regression to the real object." *Brit. J. Psychol.* XXI, 339–59 and XXII, 1–30.

(4) VERNON, M. D. (1937). *Visual Perception*. Camb. Univ. Press.

(*MS. received* 15 *December* 1937)

EXPERIMENTS ON THE CONSTANCY OF SHAPE

By W. E. MOORE

From the Psychological Laboratory, University of Manchester

I. *Introduction* (pp. 104–106).
II. *Method of experiment* (pp. 106–109).
    (1) *Aim of the experiments* (pp. 106–107).
    (2) *Apparatus* (pp. 107–109).
III. *Results of the experiments* (pp. 109–114).
IV. *Discussion* (p. 115).
V. *Summary* (pp. 115–116).
References (p. 116).

I. INTRODUCTION

It has been stated by Boring (1) that in many ways the problems of perception have always been central in systematic psychology. It is certain that the differences, or lack of correspondence, between perceptual awareness on the one hand, and either sensory stimulation or physical reality on the other, have always been the central problems of perception. And in recent years, the Gestalt psychologists have introduced these problems extensively into the experimental laboratory. A mass of data has been accumulated covering countless manifestations of this discrepancy between perceptual awareness and the pattern of sensory stimulation. One of the most readily measurable of these manifestations is that referred to as constancy of shape, or phenomenal regression to the real shape. If an object is tilted at an angle to the line of vision of the observer, the phenomenal, or seen, shape of that object is intermediate between the shape of the retinal stimulation and the real, or physical, shape of that object.

The most extensive investigations of this constancy of shape have been those of Thouless (6, 7, 8), and Brunswik's co-workers, Eissler and Klimpfinger (3, 4). All the workers have observed that the amount of constancy in any instance depends upon the presence of factors conducive to good depth perception, for example, binocular vision, good lighting, and so forth; and all the workers have observed considerable individual differences.

Thouless has come to the conclusion that the phenomenon is to be explained by the presence of spatial clues which, so to speak, "give information" as to the real shape of the object, and thus cause a compromise to be made between this and the retinal shape.

Brunswik's own interpretation (2) of the results is similar to that of Thouless, but is more extensively worked out. In the perception of a non-normally oriented object he asserts that the subject is making a compromise between tendencies to react to each of two 'poles' (to use Prof. Tolman's translation). One pole is the *real* shape of the object and the other is the *retinal* shape. This tendency to compromise between poles is regarded as general in the perceptual field, and indeed in the whole of the psychology of judgement. For example, in judging a person's intelligence, the person's actual capacity may form one pole, but other poles, such as personal appearance, may also be present, and have a considerable influence on the judgement. It is partly this generalization of his results by Brunswik that has helped to bring constancy problems directly into the focus of psychological research.

A quite different and highly elaborate theory of the constancy of shape has been built up by Koffka (5), and is largely based on one of Eissler's results, namely that the constancy appears to be different for differently shaped objects. Briefly, Koffka's explanation of the phenomena of constancy of shape rests upon two assumptions:

(1) That the frontal vertical plane is the most stable of all planes, and any surface not in that plane will tend to regress (phenomenally) towards it.

(2) Some objects have more highly organized, or more 'stable' shapes than others, and hence there will be a tendency for any object to regress to a more highly organized shape.

Suppose, then, we have a circular object, seen in section *XY* (Fig. 1) tilted from the frontal vertical plane of an observer *O*. Then, according to Koffka's explanation, there will be two sets of 'forces' acting on the ellipse that is seen at *O*.

(1) Those tending to render it phenomenally more nearly in the frontal vertical plane.

(2) Those tending to render it more circular (as the circle is a very highly organized form).

These 'forces' are referred to as *internal organizing forces*. They are both opposed by the *external organizing forces* set up by the actual distribution of retinal stimulation. The result is thus the compromise so characteristic of constancy phenomena. Both types of internal organizing forces tend to produce the circular form, the first by reducing the apparent tilt of the object, and the second by a sort of expansion of the elliptical form itself. The external forces set up by retinal stimulation tend to make the object be seen retinally, and interact with the internal forces producing the observed compromise. This explanation would account for Eissler's result concerning the different constancy for different shapes, by the fact that the internal organizing forces of the second kind would be different in the two cases, depending on the real (physical) shapes of the respective objects.

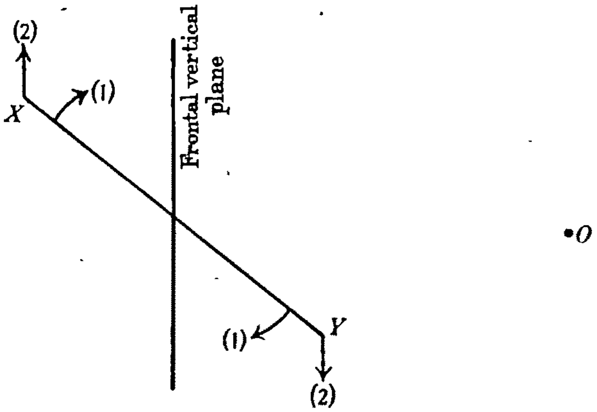


Fig. 1.

The experiments reported in this paper were designed to test this theory, the general plan being to measure the constancy of shape for two different figures; if any difference in constancy exists, it must be traced to its source, making sure that it is not due to factors such as difference in area, before ascribing it to the cause suggested by Koffka.

II. METHOD OF EXPERIMENT

(1) *Aim of the experiments*

The figures investigated were the straight line and the circle. As no particular length of line is regarded by Gestalt psychologists as more highly organized than another, whereas a circle is regarded as more highly organized than an ellipse, then if internal organizing forces do play a part, a greater degree of constancy ought to be shown by a tilted circle than a tilted line.

This was the question it was decided to test. Briefly the plan of attack was as follows:

(1) Measure the phenomenal length of a straight line, tilted from the frontal vertical plane to various angles of tilt, by comparison with shorter lines.

(2) Ditto, for a tilted circle, by comparison with ellipses.

If any difference in constancy is found, it will be necessary to perform experiments to determine whether it is actually due to the working of such organizing forces, or to other factors.

(2) *Apparatus*

The stimuli used consisted of vertical lines and ellipses. The lines were constructed of carefully straightened galvanized iron wire, soldered at right angles to horizontal lengths of the same wire which could be mounted at their ends in wooden supports *S* (Fig. 2). The vertical wires *AB* were painted white and the rest dull black.

The ellipses were constructed geometrically on white cardboard, carefully cut out, and for use were firmly fastened to the 'line' framework by means of adhesive tape. Thus the same supporting arrangement could be used for both lines and ellipses. The arrangement of the stimulus lines (or ellipses, if these were fixed to the lines) in the supporting apparatus is shown in Fig. 3. *X* and *Y* were vertical boards fastened to a heavy base. The upper stimulus line was fastened firmly into its cylindrical wooden supports, which were free to turn in circular holes in *X* and *Y*, and one of which, *S*, was long and projected through the black cloth *b* (Fig. 3) surrounding the whole apparatus, and carried a pointer *p* which indicated on a circular scale *x* the angle at which *AB* was inclined to the vertical. The lower stimulus lines *CD* were removable from their supports which were tightly inserted in *X* and *Y*.

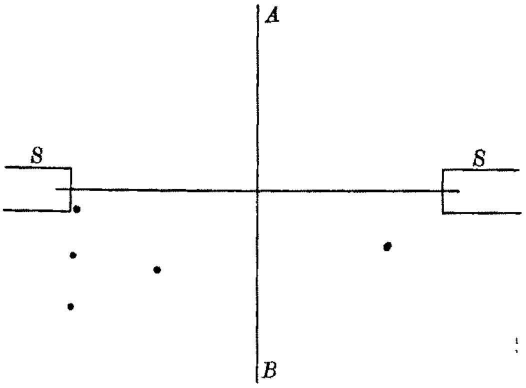


Fig. 2.

The whole of the apparatus, except the vertical stimulus wires, was painted black.

The subject sat in a low chair (Fig. 4) in front of a black screen *AB* with a chin

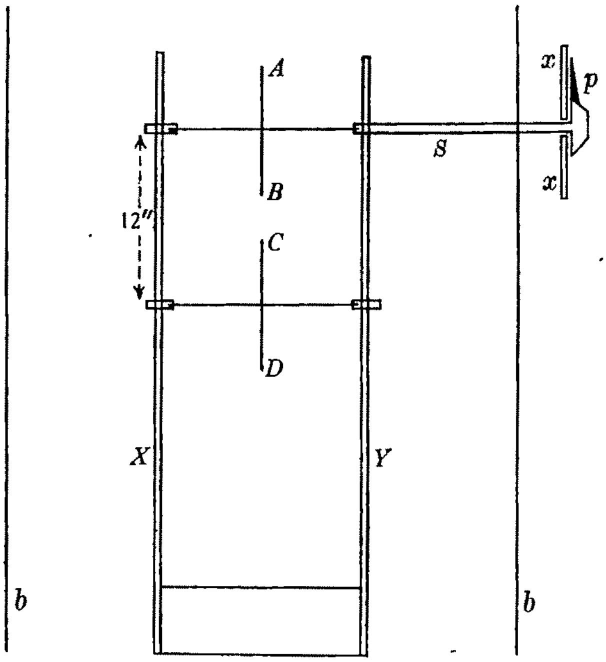


Fig. 3.

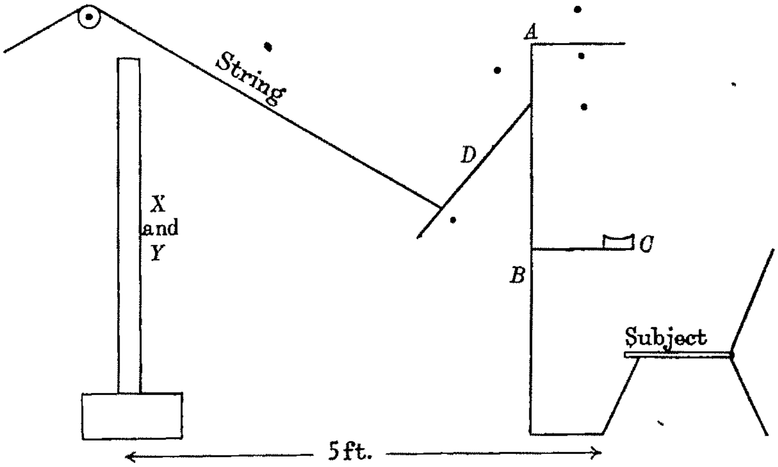


Fig. 4.

rest C and a door D, covering a rectangular aperture, through which the subject could see the stimulus lines or ellipses and their supports in a frontal vertical plane. The chin rest was at such a height from the floor as to make the eye level of the subject fall midway between the levels of the horizontal supports of the stimulus lines. The chin rest was at a distance of 5 ft. from the plane of the supports X and Y. The door D could be raised and lowered by the experimenter by means of a string. The experimenter's position during the taking of readings was, of course, close to the scale x and hidden from the subject by the cloth b.

The experimental room was darkened, and the figures were illuminated by a single floodlight placed in front of and below them, and directed towards a point midway between the centres of the two figures. The lamp was of 60 watts, and the light was diffused by passing through three thicknesses of ordinary typing paper. In this way it was found that the figures were as nearly as possible equally illuminated *for all angles of tilt of the upper figure.*

The upper stimulus line (circle) was 10 in. in length (diameter) and a series of lines less than 10 in. long and ellipses of less than 10 in. vertical axis were constructed for the lower stimulus objects.

III. RESULTS OF THE EXPERIMENTS

The subjects were six graduates and research students of psychology. The subject was instructed to report whether the upper figure appeared longer, equal to, or shorter than the lower when the screen in front of him was raised.

The method was as follows: One of the lines CD (Fig. 3) was introduced vertically in its supports. The upper line AB was rotated to any multiple of 5° between 0° and 90° in an anti-clockwise direction. The screen D was raised, and lowered as soon as the subject had made a judgement. AB was then rotated to some other position (multiple of 5°) on the scale and the judgement repeated. This process was continued until a judgement had been obtained for every 5° position on the scale. The positions were presented in random order. The above was repeated for all values of CD. After readings had been taken with a few subjects it became unnecessary to continue presenting the whole range of angles of tilt with each value of CD, and it was possible to give just five or six positions of AB around the 'equality' point.

The procedure was repeated for all subjects with both the tilted line and the tilted circle.

Table I shows a record of judgements of one subject for lines. It indicates the general method of recording judgements. L represents a judgement 'longer', S for 'shorter' and E for 'equal'. From the results as recorded in this table it was possible to read off the angle of tilt of AB necessary for it to be phenomenally equal to CD for a particular subject.

The difference between this and the angle of tilt necessary for equal retinal projection gave the measure of constancy used in these experiments. This measure will be referred to subsequently as the '*angular compensation*'. It is recorded in the final column of Table I.

Average values of this measure of constancy were then calculated for various angles of tilt for both the tilted line and the tilted circle. These values are shown in Table II (upper section). It can be seen that there is a tendency for the constancy of shape to be greater for the circle than for the line, over the explored range. This difference was deemed sufficiently large to warrant further experimentation.

Table I. *Judgement of lines by one subject*

Length of *CD* in in.	Angle of tilt										Angle of tilt for equality	Angular compensation
	0°	5°	10°	15°	20°	25°	30°	35°	40°	45°		
9·65	*L*	*L*	*L*	*E*	*E*	*S*	*S*	*S*	.	.	17·5°	2 5°
9·39	.	.	*L*	*L*	*L*	*E*	*S*	*S*	*S*	.	25·0°	5·0°
9·06	.	.	*L*	*L*	*L*	*E*	*S*	*S*	*S*	.	25·0°	0·0°
8·66	.	.	.	.	*L*	*L*	*L*	*E*	*S*	*S*	35·0°	5·0°

Table II. *Average values of angular compensation for six subjects*

Type of figure	Angular compensation for angle of tilt						
	15°	20°	25°	30°	35°	40°	45°
Line (10 in.)	0·0	1·0	2·0	3·0	3·9	—	—
Circle (10 in.)	—	7·0	8·0	9·0	10·0	10·0	—
Small circle (4 in.)	—	7·0	7·8	8·5	9·7	11·0	—
Outline circle	—	5·0	7·3	9·1	10·0	10 0	—
Prolate ellipse	—	—	9·2	9·4	9·0	9·0	—
Line (10 in.) (repeated)	2·2	2·5	3·0	3·8	3·8	—	—
Circle (10 in.) (repeated)	—	8·2	9·0	10·0	10·0	11·6	—
Small circles 'separated'	—	5·8	7·5	9·2	10·7	12·2	—
Small circles 'together'	0·8	2·0	2·9	3·7	4·4	—	—

The difference may be dependent upon a number of factors:

(1) Difference in area of the stimuli.

(2) The circle and the ellipses are surfaces, and as such may have different properties from linear figures. For example, it is often stated that differences in brightness between the upper and lower limits of a tilted figure form one of the most important clues which form the basis of constancy phenomena. In the present apparatus slight brightness differences did exist, and it is possible that these were more easily perceptible on the surface figures than on the lines.

(3) Internal organizing forces, which would give a circle a greater tendency to constancy than a straight line.

The following experiments were therefore performed to test for the presence or absence of differences in constancy due to these three factors:

(1) The constancy of shape was measured for a circle of 4 in. diameter. Exactly the same method was used.

(2) The constancy of shape was measured for an '*outline*' circle of 10 in. diameter. The stimulus objects were constructed of stiff cardboard (Fig. 5). The outline was $\frac{1}{8}$ in. wide, and the cross pieces were painted dull black. These parts were almost invisible when the figures were placed in the apparatus under experimental conditions, and the impression of 'outline' figures was successfully obtained. The method was exactly as before.

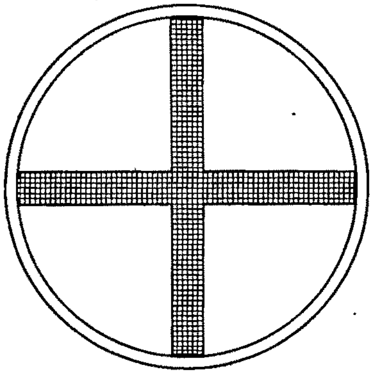


Fig. 5.

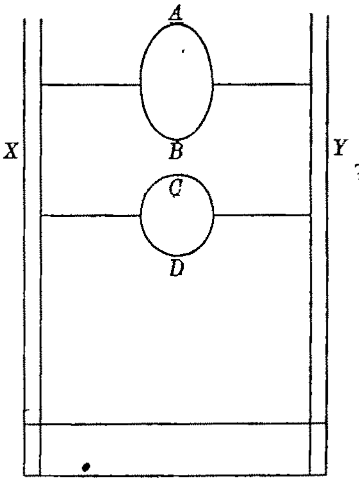


Fig. 6.

(3) The constancy of shape was measured for a series of prolate ellipses. The set-up is shown in Fig. 6. In this case the ellipse was tilted until it appeared circular. If internal organizing forces of the kind postulated by Koffka exist, then, when a prolate ellipse is tilted until it appears circular, these forces should be expected to assist in the production of the circular form; i.e. they would assist the external forces instead of opposing them, as in the case of a circle tilted to appear equal to an oblate ellipse. The constancy in the case of the prolate ellipse would, then, be expected to be less.

The results of these three experiments are shown in the second section of Table II. The values of the angular compensation are in each case not greatly different from those for the original circle. The difference originally

φ 166.54

observed between the circle and the line cannot, therefore, be ascribed to any of the three causes suggested above.

Two possibilities, however, still remain:

(1) It is possible that the results of the later experiments had been modified by practice, and that, if the subjects were re-tested for constancy with the original tilted line and circle, the difference between the two would have disappeared.

(2) It is also possible that the subjects might consciously or unconsciously be adopting different subjective methods of comparison in the two set-ups, and that this might account for the observed differences in constancy. The following experiments were therefore performed, with a view to testing these two points:

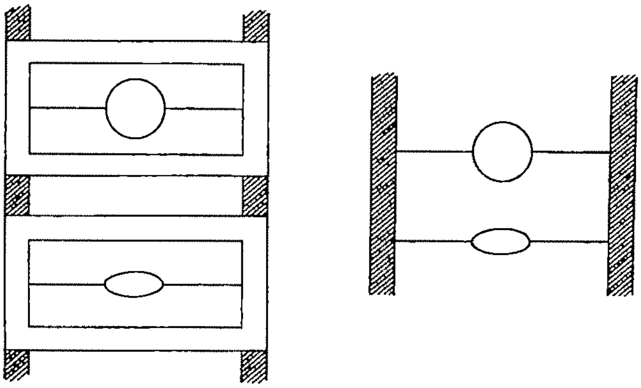


Fig. 7.

(1) The original measurements with the original line and circle were repeated.

(2) The constancy of shape for the circle of 4 in. diameter was measured in the two set-ups illustrated in Fig. 7. The first of these was designed to encourage a *separative* method of comparing the figures. The second, on the other hand, was expected to call forth a more truly simultaneous method of comparison. These set-ups were chosen more or less at random; any two arrangements likely to encourage different subjective methods of comparison would have served the same purpose.

The results of these two experiments are shown in the lower sections of Table II. From an examination of the figures it can be seen that the effect of practice has been negligible. There is a difference, in constancy, however, between the two set-ups of the second experiment ('separated' and 'together') which is comparable to that originally observed between the circle and the straight line. These results suggest, then, that a difference in subjective method did play some part in producing the original difference in constancy.

In order to confirm, or otherwise, this tentative conclusion, one further experiment was performed. The subjects were given a week's training in making the comparisons by a definitely separative method. That is to say, they were told to look first at the upper figure, then at the lower, then back at the upper, and then to make a judgement. These instructions were strictly enforced. The set-ups were the original 10 in. circle and ellipse, and the small circle arrangements of the previous experiment, as shown in Fig. 7. At the end of the week's training each subject was given four sittings with each set-up. The average angular compensation before and after the period of training for each set-up is shown in Table III.

Table III. *Average values of angular compensation before and after training in 'separative' method of comparison*

	Angular compensation for angle of tilt									
	Before training					After training				
Type of figure	20°	25°	30°	35°	40°	20°	25°	30°	35°	40°
Line (10 in.)	2·5	3·0	3·8	3·8	—	—	18·0	20·0	21·7	23·0
Circle (10 in.)	8·2	9·0	10·0	10·0	11·6	—	12·0	14·1	16·2	17·1
Circle (4 in.) 'separated'	5·8	7·5	9·2	10·7	12·2	7·0	9·2	11·6	12·2	11·0
Circle (4 in.) 'together'	2·0	2·9	3·7	4·4	—	7·2	9·2	11·4	12·2	10·0

From this table it can be seen that the week's training in using a particular subjective method had very considerably altered the constancy in some of the set-ups. The expectation was that the original differences between the circle and the line, and between the two set-ups of Fig. 7, would have been somewhat diminished. Instead of this the constancy for the line, for example, now appears to be significantly greater than that for the circle, instead of smaller, as previously. Probably the most important factor, however, is that the period of training in one subjective method had unquestionably affected the constancy in two of the set-ups.

The above experiments, and the conclusions suggested by them, have been criticized by Prof. F. C. Bartlett (in a private communication) on the grounds that the horizontal supports of the stimulus objects were not entirely invisible, and that, in reality, the tilted lines were behaving as tilted crosses. An attempt was made to meet this criticism. A new apparatus was constructed in which the horizontal supports were made of very thin piano wire, painted dull black, which was reported to be invisible by all the subjects under the experimental conditions. The constancy of shape for the original circle and the original line was inmeasured in the modified apparatus. The same subjects were used. The results are shown in Table IV. It can be seen that, although the angular compensation is in general greater than before, the circle still shows greater constancy than the line.

In the case of one subject, however, the reverse was true. His results are shown in the lower part of Table IV, and it can be seen that his constancy in these supplementary experiments is greater for the line than for the circle. In the original experiments, this subject's results followed the general trend (i.e. his constancy was greater for the circle). If the main contention of this paper is correct, a reversal of this kind in one subject might be explained by his having adopted a different subjective method or 'set' in the later experiments. On examining the introspective reports of this subject, it is found that in making the circle judgements, he relied in the first series mainly upon a comparison of the curvatures of the edges of the figures, while in the supplementary series he was concerned with comparing the actual visual heights. This may be the explanation of the reversal obtained with this subject.[1]

Table IV. *Values of angular compensation with modified apparatus*

Type of figure	Average angular compensation for five subjects for angle of tilt				
	20°	25°	30°	35°	40°
Circle	—	13·7	15·0	17·6	18·0
Line	9·0	11·0	10·0	6·8	—

Type of figure	Angular compensation for Subject F for angle of tilt						
	25°	30°	35°	40°	45°	50°	55°
Circle	20·0	21·8	23·8	29·0	32·0	35·0	37·5
Line	27·5	30·0	32·5	35·0	37·5	40·0	42·5

Apart from this subject, the results show that, in the main, a difference in constancy between the tilted circle and the tilted line, similar to that originally found, exists when the supporting wires are completely invisible.

---

[1] This subject unfortunately left England a few days after taking the later series, and the matter could not be further investigated.

IV. DISCUSSION

It is clear from the small number of subjects and the imperfections of the apparatus that any conclusions which may be drawn from the foregoing experiments must, at best, be accepted as suggestive only. The inferences which are fairly definitely indicated, however, are these:

(1) Other things being equal, a tilted circle will show a greater degree of constancy than a tilted line.

(2) This difference must not be accepted, however, as evidence for the Gestalt theory of constancy described in the introduction, for two reasons:

(*a*) A tilted prolate ellipse shows quite as much constancy as a tilted circle. This would not be the case if the Gestalt explanation were valid.

(*b*) The difference in constancy between a tilted line and a tilted circle, although not affected by practice, is clearly affected when the subjects are trained to use a special subjective method in viewing the material.

It is suggested, therefore, that the subjective method, and not internal organizing forces, is largely responsible for the difference.

All these conclusions must be taken as applicable only to the experimental conditions of this research, and only for adult human subjects with no special training in drawing or similar occupation. There is an evident need for research with other groups of subjects.

V. SUMMARY

1. The constancy of shape is appreciably greater for a tilted circle than for a tilted straight line.

2. The difference is not due to the difference in area between the two figures. Nor is it due to the possibility that outline and surface figures have different visual properties.

3. The difference cannot be explained by the difference in the two cases of the internal organizing forces, as postulated by Koffka.

4. The present experiments tend to show that the difference in constancy is to be ascribed to a difference in *subjective method*. The circle and line 'set-ups' seem to encourage respectively different methods of comparison of the standard and test stimulus objects.

5. Additional experiments were performed to meet the criticism that, as the horizontal supports of the vertical lines were visible, these figures

might be behaving as tilted crosses. When these supports were invisible to all the subjects, the circle still exhibited greater constancy than the line.

REFERENCES

(1) BORING, E. G. (1929). *A History of Experimental Psychology.* New York, Century.
(2) BRUNSWIK, E. (1934). *Wahrnehmung und Gegenstandswelt.* Leipzig.
(3) EISSLER, K. (1933). *Arch. ges. Psychol.* LXXXVIII, 487.
(4) KLIMPFINGER, S. (1933). *Arch. ges. Psychol.* LXXXVIII, 551.
(5) KOFFKA, K. (1935). *Principles of Gestalt Psychology.* London, Kegan Paul.
(6) THOULESS, R. H. (1931). *Brit. J. Psychol.* XXI, 339.
(7) —— (1931). *Brit. J. Psychol.* XXII, 1.
(8) —— (1932). *Brit. J. Psychol.* XXII, 216.

(*Manuscript received* 27 *February* 1938)

A NOTE ON AN IMPROVED DESIGN OF VOICE KEY

By D. F. VINCENT

From the Laboratory of the National Institute of Industrial Psychology

I. *Constructional details of the voice key* (p. 117–119).
II. *The relay and electric circuits* (p. 119–120).
III. *Details of adjustment* (p. 120).

I. CONSTRUCTIONAL DETAILS OF THE VOICE KEY

THE object of this paper is to describe in detail a type of sensitive voice key that does not require a thermionic amplifier and is very cheap to make. The instrument described was built in the workshop of the National Institute of Industrial Psychology for use in research upon light-coloured 'blackboards'[1]; a very similar instrument was constructed by the writer seven years ago and has been in use in the psychology laboratory of Reading University ever since.

This instrument does not differ in principle from the voice key of standard design but it differs very considerably in constructional details. The essentials of the design are shown in Figs. 1 and 2. A voice key constructed about double the dimensions indicated in Fig. 2 would represent a convenient compromise between bulk and sensitivity. The horn, soundbox and diaphragm are all soldered together into one piece. The horn is made from any sheet metal and may be as light as is compatible with rigidity. The soundbox requires to be fairly rigid and should preferably not be made of metal thinner than 24 gauge. The diaphragm must be made of very thin springy sheet metal: 'hard rolled' copper 2 mils (i.e. 0·002 in.) has been found to be very satisfactory and is readily obtainable. The construction of the jumper which is shown in greater detail in Fig. 2 is most important.

At the top is about an inch of clock spring. It is shown riveted to the brass supporting piece, but it may be attached in any way convenient provided that the temper of the spring is not destroyed in the process. The hair spring below it is soldered on, and the brass contact piece at the bottom (about a quarter of an inch of 16 gauge brass wire) is also soldered.

---

[1] Seymour, W. D. (1937). "An improved blackboard." *Brit. J. educ. Psychol.* VII, 259.

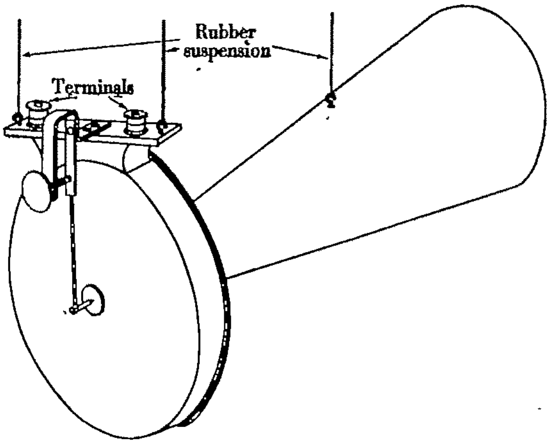


Fig. 1.

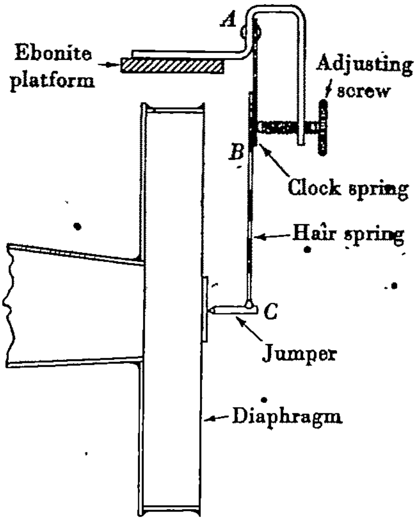


Fig. 2.

The nature of this spring is most important. Too heavy a gauge of spring will make the instrument insensitive, while a spring that is insufficiently resilient (such as a narrow strip of hard rolled copper or brass foil) will be unsatisfactory because it will continually need adjustment. A very satisfactory form of spring can be made by taking a piece of 28 gauge steel wire (a single strand from a bowden cable) and hammering it out till it forms a ribbon about four times its original width. The contact piece bears on a disc of thin sheet brass soldered to the centre of the copper diaphragm. Platinum contacts would be better in theory but 'brass to brass' works very well in practice. The contact pieces may be cleaned by rubbing gently with a small strip of very fine emery cloth when necessary.

The suspension of the instrument is also important. Although a rubber suspension employing the small rubber rings sold by stationers has always been used, there is no doubt that the helical metal springs would be equally satisfactory. Whatever suspension is used, it must allow the instrument to oscillate up and down freely and not more rapidly than one excursion per second. (The side to side 'pendulum-wise' movement will be more rapid than this.) An alternative form of suspension that makes it very easy to adjust the instrument to any angle is shown in Fig. 3.

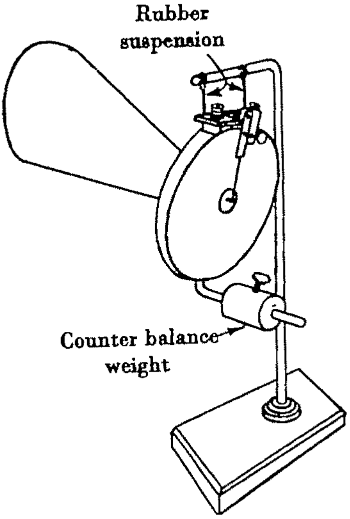


Fig. 3.

II. THE RELAY AND ELECTRIC CIRCUITS

A suitable relay and the electric circuit employed is shown in Fig. 4. No doubt there are many commercial designs of relay that could be employed satisfactorily. If a relay of this form is built it is an advantage to use silicon steel transformer stampings, such as stalloy, for the magnet core. The residual magnetization of a core of transformer stampings is negligible. This not only keeps down the time lag of the relay but reduces its variability. The resistance R_2, which may be anything from 100 to 1000 ohms, is essential to prevent sparking at the contacts of the voice key. The rheostat R_1 is used to adjust the current to the smallest value at which the relay will operate.

The relay used in the research upon light-coloured 'blackboards' had a stalloy core half an inch by a quarter of an inch in section and each bobbin was wound with about three hundred turns of 28 S.W.G. copper wire. The current through this winding and the contacts of the voice key, when the latter was set for medium sensitivity, was about a tenth of an ampère. The resistance R_2 effectively prevents all sparking at the contacts, so that the value of this current is unimportant. The magnet windings were chosen so that this part of the instrument could be wired up with the same 4 V. accumulator used with the chronograph. With a different setting of R_1 they would be quite suitable for use with a 2 V. accumulator. If it were desired to use a flash lamp battery for this circuit it would be better for reasons of economy to use a finer wire and more turns. About twice the number of turns of 32 s.w.g. would probably be satisfactory. The effect of sound waves upon this voice key is to produce a small fluctuation in a comparatively large steady current. The use of a commercial relay, which is operated by the cessation of a very small current, would necessitate a Wheatstone bridge type of circuit to balance out the steady current. It would still be essential to include the resistance R_2.

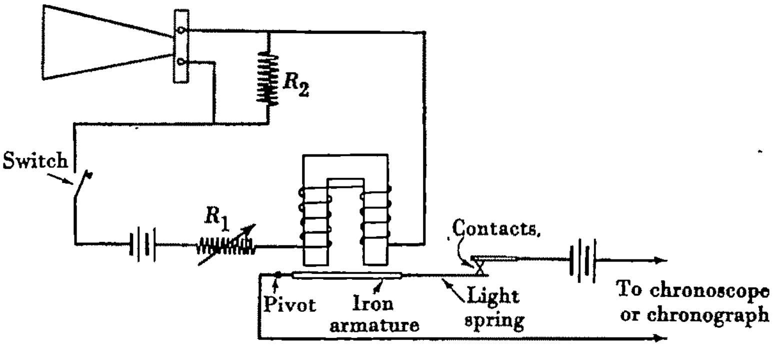


Fig. 4.

III. Details of adjustment

When the apparatus is set up for use the adjusting screw of the voice key is turned till the jumper just makes contact with the diaphragm; the rheostat R_1 is then adjusted so that it passes the smallest current that will hold up the armature. Next, the adjusting screw is rotated very slowly till the relay breaks. The adjusting screw is then turned back a very little way and the final adjustment to the right degree of sensitivity is found by trial and error. A voice key designed in this way can be so adjusted that it reacts to any sound made in a normal speaking voice, even to the sibilant consonants such as S and Z, but if it is as finely adjusted as this it may react to irrelevant sounds. It is a good plan to adjust it to fire at the vowel 'ee' or a word containing it. If 'ee' will make it fire, any other vowel will, and consequently any spoken word.

(*Manuscript received* 28 *May* 1938)

CHARACTERISTIC MOTIVATION IN THE ACTIVITIES OF SCHOOL-GIRLS

By M. D. VERNON

From the Cambridge Psychological Laboratory

Part I

I. *General description of methods of investigation* (pp. 121–124).
  (1) *Experimental method* (pp. 121–123).
  (2) *Presentation of results* (pp. 123–124).
II. *The functioning of the drives in individual cases* (pp. 124–143).

I. General description of methods of investigation

(1) *Experimental method*

The experiments here described were carried out in order to discover if it is possible to detect the existence, in girls of school age, of underlying drives or tendencies to action which determine the main lines of their behaviour, and may be expected to direct them in future towards certain occupations or ways of life. An experiment reported in a previous paper (*Brit. J. educ. Psychol.* (1937), vii, 302 and (1938), viii, 1) indicated that it seemed possible to discover such drives or behaviour tendencies in University students. The more important of these drives were there described, and their manner of functioning illustrated in the remarks made by the students in answer to a series of questions about their interests, activities, etc. Clearly it is both interesting and important to discover if such drives are operating sufficiently strongly in younger girls to have a decisive effect on their school careers and on their selection of a future career or occupation.

The following groups of girls were selected for study:

Group (a): 29 girls, aged 13½–16 (average 14–9) from a provincial High School; these girls made up the greater part of two parallel forms working for the School Certificate examination, which they were to take in rather less than two years from the time of the experiment. Their education was in no way specialized.

Group (*b*): 28 girls, aged 12–14½ (average 13–6) from a rural Senior School; they were the oldest and most intelligent girls in the school, and were for the most part shortly due to leave. Their curriculum consisted partly of theoretical subjects (English, Arithmetic, etc.) and partly of practical ones (domestic science).

Group (*c*): 51 girls, aged 15½–19 (average 17–6) from a large boarding school in the south of England; these girls were all doing post-School Certificate work. The majority of them had passed the School Certificate, though some had failed in it or not taken it. Some of these girls were specializing, with a view to entering the Universities or taking up specialized careers. The remainder were doing a general educational course including theoretical and practical subjects (e.g. domestic science) until the end of the school year, when they would leave school.

Series of questions (one for each group) were drawn up, similar in type though different in detail from those used for the University students. That is to say, the questions were directed towards finding out interests and occupational wishes; but throughout less weight was assigned to the direct answers than to the indirect evidence provided by the reasons given for those interests, the manner in which they were described, etc. As additional information about the girls there were obtained: For Groups (*a*) and (*c*), short descriptive sketches from the headmistresses of the parentage and outstanding characteristics (intellectual, temperamental and behavioural) of each girl. For Group (*b*), brief characterizations and opinions about each girl were obtained from two of the women teachers in the school. For the evaluation of school performance, some assistance was obtained for Group (*a*) from class marks; for Group (*c*), from School Certificate results, when obtained; and for Group (*b*), from intelligence tests previously administered by Dr A. Crowther.

The general form of the questions was as follows:

(1) *Group* (*a*)

Can you tell me what you most enjoy doing at school...what lessons you like best? What part of (so-and-so) do you like best?

Do you play games? Which do you like best? Are you in any of the teams? A house representative?

Do you belong to the orchestra? Choir? Debating society? Do you enjoy these? Do you like speaking at the debating society?

What sort of things do you like doing out of school...in the holidays? Which do you like best, term or holidays? What is it about term (holidays) that makes you like it best?

Do you know a lot of girls at school? Have you any special friends? Do you do things with them out of school?

What are the mistresses like...are they nice? Any of them specially?

Have you thought at all about what you want to do when you leave school? What is it about (so-and-so) that makes you want to do it? Have you ever thought of anything else you would like to do? Are your parents quite glad for you to do (so-and-so)? Are you keen to have a career, or would you rather just stay at home?

(2) *Group* (*b*)

Can you tell me what you enjoy most doing at school...what lessons you like best? What part of (so-and-so) do you like best?

What sort of things do you do when you are out of school? In the week-ends? Do you ever do things with the other girls out of school?

Which do you like best, school or holidays? What is it about school (holidays) that makes you like it best?

Have you thought at all what you want to do when you leave school? What is it about (so-and-so) that makes you want to do it? Which would you like best, town or country? Are your parents pleased you should do (so-and-so)?

Shall you be glad or sorry to leave school? What is it that makes you glad (sorry)?

(3) *Group* (*c*)

Can you tell me what you most enjoy doing at school...first of all in school hours? What part of (so-and-so) is it that you like best?

Do you do much work alone, or mostly in class? Which do you prefer?

What do you like best doing out of school hours?

Do you play games? Which do you like best? Are you in any of the teams? Captain?

Do you go in for: music—singing and/or playing—; acting and/or elocution; arts and/or craft-work; debating—do you do much speaking? Do you belong to any of the clubs?

Are you a prefect? Do the duties take up much of your time? Do you enjoy them, or are they rather a bore?

Have you many friends at school? Some particular ones? Do you see anything of them in the holidays?

What sort of things do you like doing in the holidays? Do you like being in the town or the country best? Do you ever go away for holidays with your family? Which do you like best, term or holidays? What is it about term (holidays) that makes you like it best?

Have you thought at all about what you want to do when you leave school? What is it about (so-and-so) that makes you want to do it? Have you ever thought of any other jobs you would like to do? Are your parents glad for you to do that? Are you keen on having a career, or would you rather just stay at home? Shall you be sorry to leave school or quite glad?

(2) *Presentation of results*

It is difficult to choose the best method of presenting the experimental results. Although there seemed to be no doubt as to the existence of fundamental drives such as those described in the previous paper, yet it did not seem desirable to study these drives or their manner of functioning in the abstract, apart from a general description of the individual who possessed them. It is possible to trace a general similarity in the nature and functioning of a given drive in a number of different individuals. But individual differences do exist; the drive is modified by other drives, and the behaviour in which it manifests itself by the particular personality and environment of the individual. Moreover, it is greatly preferable to study individual personality functioning as a whole and showing characteristic behaviour tendencies, rather than single and separate drives functioning in isolation. Consequently it seemed best to follow the procedure adopted in the previous paper and describe individual cases whose behaviour, wishes and interests manifested the operation of various drives. Space prohibits a description of every case in all three groups, but a number of typical cases are set forth in the ensuing section.

The former procedure was also followed in that an attempt was made to characterize individuals by selecting from their answers those statements which showed up their characteristic drives most clearly. This was less easy to do with school-girls than with University students who are not only more formed in character, but also more skilled in self-expression, though they are perhaps also more concerned to conceal any undesirable tendencies. But in the following descriptions the characterizations are where necessary filled out by (1) impressions of the girl's manner and behaviour gained during the interview, (2) in Groups (*a*) and (*b*), similar impressions gained by observation of the girls' behaviour in class, (3) statements about parentage and school behaviour made by the Staff (subjective judgements and evaluations by the Staff have, however, been excluded). The class marks, School Certificate results and test results mentioned above were also referred to in some cases. It is clear that this procedure is not consistent with an accurate checking and validation against each other of the impressions proceeding from different sources. But the general aim of the experiment was to obtain adequate character studies of the girls by the use of all available means.

II. THE FUNCTIONING OF THE DRIVES IN INDIVIDUAL CASES

In this section are described a certain number of individuals who showed most clearly in their answers how their wishes and interests were determined by one or more of the main drives or tendencies to action which appeared in the University students. These drives were towards: social conformity, humanitarianism, independence, security, activity, superiority, dominance, display and pleasure. In addition to the cases in whom one or more of these drives functioned strongly and without much hindrance, are certain cases in whom there was a conflict of drives; or else the free functioning of the drive was inhibited by some natural or environmental disability to which the individual was subject. Such cases were also fairly frequent among the University students.

It will be observed that the number of individuals chosen from each of these groups to demonstrate any particular drive varies considerably. It was in fact the case that certain drives were manifested more frequently and more clearly in some groups than in others. This seems to have been due to three factors:

(1) The variation in age, and hence in intellectual and emotional maturity, between the groups;

(2) The favourable selection for intelligence and general culture in Group (*c*) and part of Group (*a*); and the unfavourable selection for these qualities in Group (*b*);

(3) The differences in social environment.

For these reasons the drives themselves varied in strength and frequency in the three groups; and the wishes, interests and activities in which they manifested themselves varied even more. Consequently an attempt is made to describe the distribution of the drives according to groups in Part II, Section III of this paper; and to relate distribution and manifestation to the social background in Part II, Section IV.

Social conformity is demonstrated in extreme gregariousness, desire to be with people and work with them; and also in tendencies to imitate their behaviour, echo their opinions and act in such a manner as to win their approval. Generally the 'people' whose society is desired are the nearest congenial social group—with these girls, the group of school contemporaries. If for any reason the individual cannot make contact with this group, or prefers some other group, she may feel uneasy, anxious or maladjusted.

In general the typical social conformer in Group (*c*) was quite well adjusted:

C4 likes the discussions in the class on economics and philosophy,[1] but "doesn't care for any other school subjects" and doesn't like working alone; belongs to the history club and "likes listening to others" speak, but "doesn't like speaking"; is a sub-prefect and enjoys the duties; has a lot of friends. Prefers holidays to term—"sometimes misses all the girls", but is "sorry to come back—doesn't like leaving people and things at home", though "soon gets used to school routine".

S4 is specializing in languages, but "likes doing out-door things", and reading, knitting, listening to the wireless. Is a prefect—"a responsible position, but not much duties"—and likes it. Has a lot of friends—"used to have a 'crowd', but they have left now"; sees them in the holidays. Prefers holidays—"boarding school makes you appreciate home—clothes, shops, etc." Wants to go to the University and "hopes to get ideas (about jobs) from meeting people who have done interesting things". "Mother's keen (for her to have a career), but step-father wants her to stay at home and look after the house—would bore her."

---

[1] This interest was somewhat of a social tradition in this group (see Part II).

B2 likes games, gym, country dancing, housewifery, acting, art—"makes papier maché animals" and has "designed an under-water scene for a bathroom". Is a prefect—"finds it interesting"—"you have more jobs and responsibility", but she "likes doing these". Has a lot of friends and meets them in holidays; prefers holidays to term—"likes being with family"; but "loves it here", will be "sorry to leave", it "makes a change having to do what you don't want". Is going to take up nursery school work—"loves children", "likes small boys best".

In Group (*a*) H3 also presented a case of the well-adjusted social conformer. Though inferior to many of her companions in culture, intelligence and social standing, she was cheerful, friendly and contented:

She likes "doing things with her fingers"—drawing—designs or decorative work—needlework and knitting; would "much rather listen to others in debates" than talk herself; has many friends in and out of school, rides with them and meets them at Guide meetings; is very keen on riding—"spends all her free time on it"—helps her father (a riding master) with his horses; would like to take up something to do with riding, but realizes "there is no opening in it—there are too many riding schools"; but thinks of being "a child nurse or a young children's teacher"—"has had a lot to do with little children, likes them".

In the Group (*c*) girls described above there were slight indications of conflict between the home and school groups and their respective loyalties. In some of the cases in Group (*a*) there was a stronger conflict between school and home; the atmosphere of the former was academic and highly cultured, and that of the latter simple, bourgeois and domestic. V1, for instance, found it hard to adopt school interests and activities, since her own bent was towards the practical and domestic:

She is "fond of domestic science"—"does quite a lot at home" and "would like to do more at school"; "likes doing things with her hands" and "would like to take up domestic science". "Wouldn't like it all holidays"—"doesn't always know what to do at home"; "lives in a village, not much to do there". Has several friends at school but they don't live near her; has some friends at home, "does things with them".

E1 found some difficulty in adopting the interests, etc., of her present school surroundings, and longed for the familiar atmosphere of the school she had just left:

She likes Guides—"it used to be fun in Scotland" but she's "new to it here"; knows some of the girls in her class and is friends with one who is also a Guide; hopes to go to Scotland next summer and see her friends there; likes writing letters to them. Likes "playing with small children, but doesn't know any here". Has "thought of thousands of things she might do" as a career, but doesn't know which to choose; "would like to decide soon"; would "like to do something with horses, dogs or small children"; "someone suggested a hospital almoner—sounds interesting—would enjoy that".

E1 showed a tendency to avoid the difficulty in conforming to a social setting which was unfamiliar and possibly too intelligent and highly academic, by an escape into the society of children or animals. K1, also from Group (*a*), showed a still more marked example of the same tendency. She was very shy in manner, found the use of words difficult; she was out of touch with people, and the society of animals was highly congenial to her:

She is not friends with many girls, and hasn't any friends at home. Prefers holidays to term—"you can do what you like"; likes riding, "by herself or with a groom or farmer"; likes walking; keeps pets—"a cat and twelve bantams"—; likes working with animals and poultry and "would like to have a farm"; her "father has bought a farm and will keep it for her".

M5 in Group (*c*) presented rather a similar case. She was plain, awkward and unintelligent, and nervous and fidgety in manner. She seemed to be suffering from a strong feeling of inferiority which appeared both in her desire to escape into the society of animals, and in a compensatory eagerness to be 'presented' at Court:

She likes best playing the piano and organ, but prefers holidays to school—"likes home better". Likes riding best in the holidays. Is leaving at Easter and being 'presented' at Court—"is longing for it—has heard so much about it"—it is "an experience you never forget". Wants to do something when she leaves; there are "lots of things she'd like to do, but is not clever enough"; would like to do something with animals—"kennel or stable work"; "wanted to be a vet., but is not clever enough".

In Group (*b*) the majority of girls seemed to be strongly motivated by social conformity, and to experience little conflict or difficulty in achieving that conformity. It is not, unfortunately, very easy to illustrate this from their answers to questions, since inarticulateness and inability to express themselves, especially to strangers, formed part of their customary social behaviour. R1 affords as good an example as any of a large number of similar cases:

She likes physical training, music, needlework and domestic science, especially cooking and baking; out of school likes knitting and reading schoolgirl tales; at week-ends takes a baby out. Prefers school to holidays—"there is nothing to do in holidays"—"nowhere to go"—but "plenty to do at school". When she leaves she "wants to go in for a nursemaid's job" or a housemaid's; "Mother thinks a nursemaid is a better job and she is fond of children". "Likes staying on at school but won't mind leaving"; "Mother says she might as well stay on".

Here the wishes of parents were in harmony with the wishes of the child and with the activities customary in this group. In the case of H1, however, the parents (her father was a policeman) wished the child to rise above the social level of the group and to enter an occupation of superior social status for which the child was not really fitted. She was very docile in nature, eager and anxious to conform both to her parents' wishes and to the usual activities of the group, but found it hard to do so:

> She likes physical training, science, music and needlework; out of school she spends a lot of time "doing lessons for a scholarship"; "helps mother", "sometimes goes out in the car", but "doesn't do things with friends". Prefers school to holidays —"gets tired of holidays". Is going to the Technical School (if she gets a scholarship there, which is unlikely) and wants to be "either in an office or a secretary"; or do dressmaking—would prefer this, she "likes making things".

There were also cases in Groups (*a*) and (*c*) of girls who were being forced along by the tendency to conform to their parents' ambitious wishes, or were worried and anxious because they felt unable to do so. An example of the former type was A1 in Class (*a*). She had little natural ambition, but her father was eager for her to succeed in an academic career:

> She likes maths. because it is "logical and sensible; likes facts, not imaginary things". Out of school likes reading—"chiefly adventure stories", but has "not much time out of school". "Father wants her to take up maths."; she "wants to go to the University"; "should like to do something with maths. but not teach it; everyone seems to think she ought to teach, but she doesn't fancy it"; "would dislike trying to make other people understand".

H5 in Group (*c*), a girl of some intelligence and independence of mind, was nevertheless oppressed by the difficulties of living up to the standard of attainment set by school and parents, and of conforming to the wishes of both these groups; and by the feeling of moral obligation to serve the community in general:

> She likes history; loves reading; likes working alone—"you don't have to worry with other people's difficulties if working alone". Likes both school and holidays, but "doesn't like the change from one to the other", it's "such a complete change of people"; in the holidays likes "walking for miles and miles in the country among the trees and hills" and doesn't like the town—"you have to wear hats, etc., in the town". Is working for the University but "may not actually go there"—"it costs a lot and there are younger ones to be educated". Her "parents thought of making her do secretarial work", but she's "not thrilled by it"—"has never met anything yet which thrilled her, but feels there must be something round the corner"; "might possibly do social work, but this might be very depressing"; though "somebody's got to do that

kind of thing, people are always wanted for it". "Wants to make her mark at something"; it would be "quite amusing to stay at home, but she wouldn't really like to do nothing".

Most pathetic of all was F1 in Group (*a*), whose father was a domestic tyrant; he sent her to a secondary school, for which she was quite unfitted; and she was cheating in her school work in order to keep up to a standard which she could not reach. She wanted to take up child nursing, but he was forcing her to take a secretarial training:

She doesn't play games; but out of school "likes roaming round the country finding flowers"; likes gardening. Likes term better than holidays because she "meets more people"; likes "looking after children whenever she gets the chance"; wants to look after children when she leaves—"has always wanted to ever since she can remember"; her sister is matron in a Waifs and Strays Home and she would like to work there—"anywhere where she could help"; "nice there because the children are happy"; would be hard work, but "there is hard work in everything"; mother approves of this, but "father wants her to do shorthand and typing"; is leaving at Christmas, and will do shorthand and typing for a bit.

Humanitarianism. This term covers what was in fact a complex of tendencies towards the furthering of interest in, or concern for, people as such—as human beings living and functioning in their own right—not regarded simply as a more or less pleasurable part of the social environment. Sometimes, but not always, it included tendencies towards sympathy, service, help and protection. It was to be distinguished from the social conformity drive which included (1) a tendency towards being with people, obtaining social contacts, working with people; and also in some cases (2) a moral obligation towards social service of some kind, dictated by the desire to win social approval for activities which are regarded as highly admirable in women. Thus we saw that quite a number of the girls among the social conformers (notably H3 (see p. 126)) wished to take up work involving social service or the care and protection of young children. But it was not possible to say definitely if they were motivated by a genuine and inherent humanitarian drive, or by the wish to undertake a socially approved and admired occupation.

It was, indeed, a matter of some difficulty to select cases of genuine humanitarianism. B3 in Group (*c*) appeared to be such a case. She was American and had come to the school as an 'exchange' from an American school; and it is possible that even she was to some extent conforming to the kindly and friendly customs and traditions of that social background:

She is interested in English literature and history—not modern post-war history, but "something you can really see, which you can't with your own period"; in music, art. "Doesn't like sports, except individual ones"; "doesn't enjoy the game for itself, but likes being with the people she likes". Likes working alone or with three or four friends, but also "likes discussing things" in class; had many special friends in her other, American, school, but English girls "are hard to make friends with at first". In holidays stays with an English family, and "likes doing things with the family"; also "likes meeting people". Her school in America was a long way from home, it was "more interesting to be with different people in a different part of the country". When she has finished at the University, will do "anything that seems interesting", possibly teaching; is "very fond of children". "Would marry if she fell in love—has no anxiety for a career as such, but would like to do things she's interested in; also one ought to work for a bit after leaving the University, before getting married."

G2, a girl of very high intelligence, seemed to be a case of genuine humanitarianism, functioning as a desire to do social service, which conflicted with her ambition to do well in academic pursuits. But her humanitarianism may have been reinforced by the influence of her father, whose feeling of social obligation and responsibility was very strong:

She is most interested in modern languages, but also likes writing, acting, art. Wants to go to the University; "if she could get a good post in University work, would like one, in modern languages or literature"; if not, would like "social service of some kind as a career to fall back on—would like to do probation work"; "would like to do something useful"; academic life is "nice as a background, but unless a very exciting job offered, would prefer a more material (? practical) type of job".

C3 was interested in people in general and in their activities, and also had a considerable leaning towards social service, with possibly a somewhat uneasy feeling of social obligation:

She likes art best, and intends to take up architecture, because she "wanted to do something in business and in art"; intends to go to a school of architecture and train, then "start in an office"; "would like to do town planning and re-conditioning". But she dislikes the physics she has to learn in connexion with architecture—it is "inhuman". Likes history—"is interested in the past", especially the Italian Renaissance—they were "more civilized than the English". Likes playing netball, but is "not a sportswoman"; "hates the seriousness of it all"; prefers "reading history and historical novels" to most games. Is a sub-prefect, but hasn't many duties—it is "chiefly being good" which "isn't too much of a strain"; "very much enjoys school—thinks of 'life' more at school than in the holidays"; will be sorry to leave school—"there's a certain stability about it—going out into the wide world is rather frightening" though she will be "happy once she's made the break". If she fails in her architectural work, "might go into politics, but doesn't want to frightfully"; "might like a parliamentary career, but mightn't be convincing as a speaker"; but "would like to be a J.P. in a children's court".

In D2 the interest in people was more detached, with less feeling of sympathy, and no sense of obligation:

She is specializing in languages; likes all kinds of literature but "is also interested in people and in foreign civilizations". "Enjoys being a prefect"—"is interested in people—likes watching them and seeing how they develope"; was going to try Froebel work, but someone suggested she might go to the University, so she thought she would try it; had "thought of journalism after the University—something with travelling" because she is "interested in international relations"; also "had thought of teaching" —would like "studying the people she taught, watching their reactions, but would get bored with teaching the same thing over and over again".

In T2 there probably were genuine humanitarian tendencies. But they were distorted and overlaid by loyalties and obligations arising from conformity to the wishes and ideas of parents, school staff, school companions, etc. In consequence she was confused, anxious and undecided:

She likes work best of school activities; is specializing in modern languages and likes the literature, but is "also interested in the countries themselves"; likes economics and "is interested in international politics"; "is trying to learn to speak" in political debates—"does quite a lot, though she doesn't like it much"; also "finds expressing herself in writing difficult". Is a prefect and the second head of her house; finds the duties very interesting—"the responsibility in choosing other prefects"; "enjoys routine work, remembering things, also keeping discipline—though not ticking people off". Also likes seeing a lot of the housemistresses—the young ones "understand your point of view better than the older ones and are easier to talk to"; there is "a crowd of varied people on the staff, such interesting people to talk to, with outside interests". Had meant to go to College, but "doesn't much want to go now"; would like "to be a travelling secretary to some one who goes abroad"; "got so interested in countries from travelling, but doesn't know how to link up her ideas about them"; would also like to do some social work, at least "as an aside", for one or two evenings a week; "lives in an atmosphere of social work at home"—her father is a mayor—and also "is interested in people as such". Has a "very happy home"—"has an absolutely free choice" as to what she will do.

In some cases there was a conflict between independence on the one hand (with possibly some superiority) and humanitarianism on the other. This appeared in F2, a German Jewish girl who had been sent to school in England to escape conditions in Germany:

She likes best music, English literature, and history—particularly Renaissance history; likes speaking in political debates—"likes to find out what different people think even if it is against her"; is a sub-prefect—one "comes more into contact with the other girls" than the prefects do; "doesn't like it when someone is too bossy"—"otherwise is a good system". Has got to know a lot of people in the school although they "seem to keep in sets rather"; will be sorry to leave—"you meet so many different people" at school; but "will be glad of a change too, and of a free-er atmosphere"—"everything goes to bells here". Doesn't know how to decide what to do when she leaves—might do some work in languages, might teach English but "wouldn't like it much—you just stay a teacher", and "hasn't enough patience"; is also "interested in nursing" but there are "many things against it".

L1 in Group (*a*) was an example of rather a similar conflict. But she felt the impossibility of taking up the career she desired as a threat to her independence, which was thus suffering from a rather painful inhibition:

> She likes doing practical science, "seeing the actual thing and doing it for yourself"; but also likes history, especially modern history "because it's about people" and geography "because it's about where people live and how the environment affects them and their trade". Gets "worried with school work, so much of it seems so futile", "so many petty things to be done at school"; "school life is tiresome", the girls in her form are much younger than she is and are therefore "hard to make friends with". Wanted to be a doctor, and as she has been "thwarted" in that (her health was not good enough), is not very keen on a regular job; but might take up singing and concert work—wouldn't like to teach music, "hasn't enough patience".

W2, also in Group (*a*), was two years younger than F2 and L1; and probably for that reason the humanitarian tendencies were less clearly visible. But her occupational wishes suggested an underlying humanitarianism; her reason for not wanting to be a nurse or a doctor showed a rather exaggerated feeling of social obligation, as well as a drive towards independence:

> She likes reading history and English literature and writing essays, doing craftwork and drawing—likes painting "big things, in poster paints, not designs so much as pictures—has been doing illustrations for 'Job'"; likes singing, playing in the orchestra; likes going on the river and "exploring how far she can get", but regrets that so few of her friends "are allowed to go out alone" on the river. Has thought of being a missionary, a policewoman, a teacher—teaching younger children, not older ones because she "wouldn't know what they think about"; but not a secretary, because she "hates sitting tied up indoors all day long"; or a nurse or a doctor, because she would "hate to think people's lives depended on her".

Independence. The above three cases all gave some idea of the functioning of the independence drive in directing the individual towards independent activity, original interests, occupations requiring initiative and the ability to stand alone. Perhaps the clearest case of this drive was that of M1, a girl in Group (*c*) of very high intelligence and maturity of outlook, who had travelled a good deal and lived abroad. She also showed some tendency towards superiority, but this was tempered by a half-humorous acceptance of other people and their peculiarities:

> She is most interested in modern languages—"does them chiefly for their literature"; but also likes being on the Continent—"likes the people, likes speaking French". Used to act, but has not time now—"enjoys watching others more"; likes tennis—is not particularly good at it but "gets fun from it"; doesn't belong to any of the school clubs—one "can get as much out of reading for oneself as from going to the clubs". Is a sub-prefect; it is a "terrible bore—no one likes it, but they all hope to be a prefect and get a study". Prefers holidays to school—"likes the rest from school routine; gets very boring seeing the same people day by day and also knowing what's going to

happen each moment". Wants to go to the University; and "would like to get a well-paid job, with liberty to do and think what she likes, and some leisure—something not completely unsympathetic"; possibly the Civil Service, or "free-lance journalism"; "has thought of secretarial work—interesting if one got someone interesting to work for". Doesn't want to do teaching—"hasn't got the patience".

Another example in Group (*c*) was that of K 1, who was, however, much less intelligent in her outlook:

She is most interested in history, especially political history, and likes the Tudor and Stuart periods best; runs the political debates, "speaks a lot and organizes too". Is "quite glad not to be a prefect, the duties are rather a bore"; prefers holidays to term—"you can do what you like" (in holidays); "gets rather sick of school". Is going to the University, but "can't think what she'll do afterwards"; "would like something in the history line"; "has no patience with children, to teach them"; her mother is "keen on secretarial work", but she thinks it would be "rather dull", but might "take it up if she can find an interesting opening". Might go in for army nursing, though "perhaps she is not suited for it"; but it would give an opportunity to go abroad—"would like to go abroad".

In the case of H 2 in Group (*a*) we observe what occurs when natural independence is thwarted, in this case by the somewhat brusque and autocratic treatment meted out by her father. She was very offhand in manner and aggressive in her attitude towards authority:

She prefers holidays to school because they give her more freedom; is very fond of drawing—likes drawing "anything out of her head", but mostly does it out of school—does a lot when she has time; also likes reading, chiefly travel books. When she leaves school would like to do something but "doesn't want to do a regular piece of work"; knows "lots of things she wouldn't like to do"—being a nurse or a schoolmistress; it would be "awful to stand up and talk to people who are not interested"—"awful to be treated as the girls at her school treat some of the mistresses".

Security. The drive towards security was perhaps antithetical to the drive towards independence, since it directed the individual towards a safe, familiar, reliable occupation with little risk attached, and generally little individual initiative required. It was infrequent among the Group (*c*) girls, who were usually fairly well assured of financial and social security. But one or two cases occurred in girls whose parents were in rather a precarious financial position:

Thus M 2 wants to write, but "writing is not paying enough to start on—wants something definite and solid"; "will have to get a job of some sort"; may "probably teach first" but won't like it much.

H 1 "would like to travel about for a firm and interview people in business"—but it is "difficult to get a job". "May do the Civil Service (Executive)—her father wants her to, thinks it a safer job, you are sure of it once you get in." She "has always really wanted to go on the stage, but is afraid she has not got the talent"; "may get a year at the R.A.D.A."; it is "rather precarious—she wants something to fall back on if she can't get on on the stage".

W6 "had thought of the Civil Service—thinks she might like it—steady income, safe job, but difficult to get into".

One might suppose that economic insecurity would press much more hardly on the girls in Groups (*a*) and (*b*), since almost all of them were far less well-to-do. Yet curiously enough there was no reference in Group (*b*) to getting "a safe job". One or two girls in Group (*a*) alluded to it, though actually these did not come from the poorest homes; and it seemed that what they really desired was to obtain a recognized position in society such as the wage-earner may hope for:

H1 "keeps changing her mind about jobs"; "first thought of teaching"—but it would be "a long while to wait before earning anything", "doesn't want to wait till she's over twenty" before she begins to earn anything; would like the Civil Service or "some sort of secretarial job—an office, not a private secretary".

H4 "used to want to be a school-mistress—would hate it now—so difficult, would never know what to say or how to answer questions"; it would be "rather nice to do something settled—wouldn't mind it being hard work if she liked it".

It seems probable that for some of the girls in Groups (*a*) and (*b*) security consisted in the care and support of their parents. The naturally timid child was more aware of the fear of separation from parents than of the fear of poverty or unemployment. Particularly was this the case with certain girls who were retarded in their emotional development. This was found among some of the girls in Group (*b*) who seemed to be very childish for their ages. B1 in particular was very shy, nervous and fidgety, easily distracted by interruptions, and finding it very hard to answer questions:

She prefers holidays to school, because "she can have Mother then". Wants to go into service "in the country, near home"; "sometimes wishes she could always be at school", though she "wants to work".

B1 in Group (*a*) desired both economic and parental security. Her desire to cling to her parents resulted from the strong and sentimental influence exerted by her mother. B1's 'phantasy' interests showed a possible unconscious desire to escape:

These 'phantasy' interests include reading, especially sea and adventure stories, going to the theatre, listening to the wireless, watching the ships at sea; also she "would love to travel". "Wants to do something when she leaves school, but doesn't yet know what"—"won't be a nurse—hates it—has always disliked it"; has thought of being a librarian or a school secretary, or being in the Civil Service, or in a bank—

"bank is a good job"; "would rather have a job where she could stay at home—would hate going away from home—has only once been away from both parents".

There was a similar case in Group (c) of P 1, a girl of nineteen, who had lived at home until she was seventeen and had then been sent to a boarding school. Her one desire was to be at home again:

> She has only been at this school a year and has "not made many friends". Prefers holidays to term—"hates being away from home"; "hated leaving" her day school and coming here; in holidays "chiefly goes out with parents"; "hates being in a town". Has decided to "live at home" when she leaves and "go into her father's factory and learn to do shoe designing"; has "got to do something, and doesn't want to go away from home to do it"; "Mother wants her to stay at home too".

Activity. There was a considerable number of individuals who seemed to be so overflowing with energy that they never ceased in the pursuit of some interest or the carrying out of some occupation. This overflow of energy we have attributed to the functioning of the activity drive. It appeared to actuate a large number of interests and occupations of very varying character, and was not usually confined to one particular form of interest or activity. The individuals in whom the drive functioned strongly were not those of the purely superficial type who readily becomes bored and flutters restlessly from one amusement to another. At the same time, it is doubtful if they ever engaged in one activity for long enough to master it thoroughly or attain first class skill in it. This drive was characteristic of many of the girls in Group (c), three of whom are described below. First, W 3 possessed much vitality and social charm, but not much originality or independence of thought:

> She "enjoys doing everything"; loves acting, dancing, games, listening to music, making pottery; likes zoology and botany, philosophy and economics, and politics especially; is head of the school and "loves the work"; is captain of the second lacrosse team—selects and coaches the team. In the holidays likes riding, yachting, walking; enjoys both term and holidays, and is "never sure which kind of holiday she likes best". Is going as an 'exchange' to an American school for a year. Would "like to get into Parliament" eventually—her "father is very keen on it"; she might go to the London School of Economics, and "do practical social work as well as theoretical"; "doesn't know if she prefers intellectual work (to practical)—is terribly interested in what goes on in the world"; or might "go to a dramatic school".

W 1 has less social charm, but more stability and perseverance:

> She likes games and dancing best, and in the holidays prefers exercise, even walking, to cinemas, etc. Had intended to take up physical training, but was not tall enough to get into a Physical Training College. She is now specializing in science—likes "human anatomy" and dissection—"likes doing things with her hands"; also

likes drawing, modelling and domestic science; is going to take up massage and electro-therapy; will do a two-year course at the Swedish Institute, and after that do a course at a hospital or go abroad to get extra training in electro-therapy.

In M4 the activity drive was reinforced by social conformity, and the conviction that she would obtain much social approval and admiration for her intellectual and cultural abilities, her moral excellence and her social sympathies:

She is specializing in English literature, likes poetry particularly; also likes writing; hopes to read English at the University. Is very fond of music, sings, plays the piano, organ and violin, and composes songs, but doesn't want to specialize in music and do it at the University, "would rather keep it as a recreation"; "musicians seem to get into such a rut, get absorbed in it and don't see anything else". Plays a prominent part in political debates, and helps organize them; went to a summer school in Geneva last year, is going again this year; "enjoyed the lectures, enjoyed it all greatly". Is very keen on games and is lacrosse captain. Is head of her house and does a lot of organizing which she enjoys; is also "responsible for the atmosphere of the house" and "has to keep discipline". Loves "being among the people at school"; really prefers term to holidays because "there's more going on"; "it's more fun to have plenty to do". Wants to teach English because "one meets interesting people doing the same thing" and she "will like it as an experience". "Ultimately would like to do Borstal work"; has heard lectures about it and is "interested in that kind of work—likes studying different kinds of people".

In C1 in Group (*a*) a strong drive towards activity was also allied with strong social conformity which in her case took the form of gregariousness —the desire to be in a group of people and work with them, and not be left alone:

She likes games and drill best—"looks forward to drill, it cheers up the morning". Also likes science, physics especially—likes learning it practically though her "experiments never come out right"; likes maths., especially doing problems; likes history, especially economic history, because it has "hard and fast rules", "reasons and results". Doesn't want a monotonous job, "must have an interesting one"; thought of being a gym. mistress, "it seemed so easy"; but might not like teaching, "it might be awful" —she "would get so cross if they were all bored"; thought of being a masseuse, might like that, but it would be "rather grim having to be nice to people at nine in the morning"; but "likes a job with a lot of people"—would hate to be a secretary because it would be too lonely.

It was difficult to pick out good examples of the activity drive in Group (*b*), since their interests were so much less varied and widespread than those of the other groups. At the same time, we mentioned that R1 preferred school to holidays because "there is nothing to do in holidays" but "plenty to do at school". A number of other girls echoed this opinion: "school occupies your mind"; "there's always something to do"; one "knows what to do all day"; whereas "there's not enough to do out of school", "you get tired of holidays unless you go away". Thus where the activity drive is strong, and intelligence and independence somewhat lacking there is liable to be some discontent and boredom.

Superiority. We have already described one or two cases (G2, F2 and M1 in Group (*c*)) who showed a certain amount of superiority allied with independence. The superiority drive directed individuals to attempt the achievement of superior intellectual or social status, or to excel others in whatever activity or occupation they engaged. It was frequent in Group (*c*), the girls in which were accustomed to possess superior culture and material possessions, and had rarely if ever experienced inferiority in social status or attainments; thus they were often inclined to regard the maintenance of their superiority as a natural right.

D1 is going to be a chartered accountant; "wanted to have some aim, not to do nothing"; "Mother wanted her to stay at home and do half-time work, but she wanted to do something more ambitious"; "wanted to take up music, but was afraid it would spoil it to work too hard at it, instead of using it as a pastime".

B4 "has always wanted to go to College since she was small"; "would like to get a research job afterwards"; "once thought of acting, doesn't think she's good enough".

S2 was doing modern languages but "thought they would only lead to teaching" so changed to geography—but hasn't yet found that geography "leads to any definite job". "Used to dislike school intensely, hated being disciplined—has got used to suffering it." Should "hate to stay on after the summer term and be the oldest in the house".

S1 took up science because she was "better at it and more interested in it than anything else". "Doesn't want to practice (medicine) but more jobs are open to a woman with a medical degree; would prefer a research job." "Will be glad to leave, would have liked to leave earlier."

R1 likes reading a lot, and seeing plays and films "if they are serious ones". Parents didn't want her to go to an English University, so she is going to an American one. "Doesn't have to work (for her living), but wants to do something—would like writing"; "has tried writing all sorts of things", but is "not satisfied with any of them". "Wanted to act but doesn't think she would be good enough."

B1 is very keen on elocution and acting, particularly the former, because she "has more faith in herself" when doing it and "gets rather nervous when acting"; there is "nothing else she likes at all as much", though she quite likes English, especially poetry, and dancing. Will be "content to leave school—will have got all she could out of it and given it all she could". Hopes to go to a dramatic college; "they develope you in every way and you can decide which line you prefer"; "once thought of taking up defective speech training", but "wouldn't like to teach in a school"—it is a "very narrow world"—"schoolmistresses become very narrow indeed and hardly ever get married". Would like the stage, but rather fears it because she mightn't be good enough—it is "rather shattering to the nerves to start with".

These answers showed up a frequent characteristic of those who were very eager to assert or maintain superiority, namely that they often desired to take up occupations such as writing, music, acting, in which individual and personal superiority can readily appear; but they were so afraid of being unable to excel sufficiently to satisfy their desire for superiority as to reject the occupation or their own efforts in it.

In Group (*a*) B3 showed an example of superiority allied with strong activity and a fair amount of independence:

> She "has an idea for three weeks" what job she would like to do; then "it passes off and she has no idea". "Ages ago wanted to be a Member of Parliament because it would be rather grand." "Would like something in which she used her brain"; might enjoy reading classics at the University, but "might get tired of it". Wouldn't "care to be a secretary—that would be dull". Doesn't want to be a doctor or a dentist—one "would be sympathizing with people all the time".

B2 in Group (*a*) was in some ways similar to R1 in Group (*c*), but showed also signs of wishing to assert her superiority through domination of others:

> She "enjoys all the school lessons"; likes reading, drawing "all sorts of things"; likes acting, though there isn't much opportunity for it in school—but she and a friend "got up a play" in the holidays; likes music, but not playing the piano—she "doesn't have proper lessons, just does it at home, and it is rather difficult to learn things properly"; "feels a strain at school"—she "spends so much energy at school, she doesn't feel like doing anything at home". Is "always changing her mind" about what she wants to be; once "wanted to be an artist, but it has worn off, she isn't good enough"; would "like to write", but when she starts to do it she "soon gets fed up with it, it gets so dull"; might like secretarial work—she is secretary to the Debating club and enjoys it—"likes arranging things in order, keeping them neat".

It is interesting to note B2's "feeling of strain" in keeping up to her own standard of achievement. T2 reported a similar feeling; but here the standard of achievement was based upon her father's abilities—he was a University lecturer in English. G1 in Group (*c*) presented a similar case; her father was a distinguished classicist. Both these girls were almost painfully shy and awkward in manner, as if they were afraid lest they should not appear as superior as they wished. Hence there were constant hesitations and withdrawals, as if they were defending themselves against possible disparagement or criticism:

> B2 (Group (*a*)) is glad by the end of term that it's holidays, because she "feels she would like a rest"—hasn't much time for doing things out of school, in term time; but is also glad when term begins, because she "wants more work"—she likes doing the work at school. She wants to go to the University. Thinks it would be nice to teach

classics, though doesn't know if she would teach them well; would enjoy "finding out how people make their mistakes, trying to make them interested".

G1 (Group (*b*)) is especially interested in classics and also "likes reading English literature, especially poetry"; there are "no other things she specially likes at school"; though she belongs to the League of Nations club she "doesn't like speaking in it". Likes "working by herself on the whole"; only knows a few people at school really well. Prefers holidays to term—one has "more time to oneself". Wants to go to the University but "isn't sure if she will be able to"; has thought of the Civil Service, but "doesn't know what posts there are"; has "not thought of anything else—it seems a long way ahead".

H5 in Group (*a*) possessed a superiority drive of a more aggressive type. She was herself deficient in natural abilities, and resented her inferiority to her two abler sisters, one older and one younger than herself. She reacted to this painful situation by an attempt to 'go out of the field', in Lewin's terminology (*A Dynamic Theory of Personality*); firstly by going abroad, secondly by earning her own living:

She "doesn't enjoy anything at school" except perhaps drawing and playing games; but "would like to be able to speak languages"—would like to be able to speak French and German very well; "hates school, hates it all, gets tired and fed up with it—would like to leave and go abroad to school"; "doesn't like anything much out of school" except "going about and seeing places". "Mummy wants her to do domestic science and then live at home, not get a job, but would get tired of that, would get fed up with home"—there's "nothing to do"; would "like to earn some money of her own"; would like to train with a friend for child nursing—would "like to be with children and teach them languages".

Cases of superiority were few among the Group (*b*) girls; most of them were quite contented to stay as they were. But since the general level of achievement was low, a girl of moderate intelligence and somewhat superior social status such as B3 (her father was a police sergeant) had little difficulty in surpassing it; and was thus more contented and self-assured than many of the girls hitherto described in this category:

She likes maths. best—"any sort of maths."; is not fond of cooking. Prefers school to holidays—"gives one something to do"—there's "not enough to do in holidays except when you go away"; "last summer they went to Folkestone and Gorlestone—she likes the sea". Wants to go to the Technical School; thinks she "will have no difficulty in this—was in top class at the Elementary School when she was only twelve"; wants to do shorthand and typing but "doesn't want a job in an office—would like a more varied job"—like the school secretary in her present school.

F1 was also of fairly good intelligence, but her father was a farm labourer, and thus she probably received less encouragement in her ambitions—possibly even some opposition. She was aggressive and sulky in manner:

> She also likes arithmetic and English—reading "adventures" particularly; prefers school to holidays—"there is never anything to do at home" (she seemed to be almost the only girl in this group who didn't help with the housework); is taking the Technical School exam., which is "not hard to get through", and "wants to work in an office".

It is interesting to note the attitude of these two girls to 'cooking', the favourite activity of so many of the girls in Group (*b*).

Dominance. Related in some cases to the superiority drive was the drive towards dominance, manifested in the directing, organizing and manipulating of other people. Such was B2 in Group (*a*) (p. 138), who was secretary to the debating club and enjoyed it—"likes arranging things in order and keeping them neat". The supposition is that it was not only things but also other people whom she liked to arrange. B4 in Group (*c*) (p. 137) was secretary to the history club, ran the League of Nations club and was captain of the netball team; and in her manner and way of speaking showed distinctly domineering tendencies. Again M4's superabundant energy showed more than a little tendency to be employed in the manipulation and organization of others, though this was usually moralized as social service. Note for instance that she 'is head of her house and does a lot of organizing which she enjoys; is also "responsible for the atmosphere of the house" and has to "keep discipline"' (p. 136). Another rather similar case was that of A2, in whom the desire to excel and to dominate was strong, but who moralized it into an exaggerated sense of duty and of the gratitude she owed to the school and to her parents:

> She likes acting and reciting best—"prefers women's parts to men's, though is often given men's ones"; is specializing in modern languages—"took them because she was best at them, but really likes English best". Runs the political debating society; is a house prefect, and "sometimes has the duties of the head of the house"; finds it "easy to keep other people in order, but not easy to keep herself in order". Has a lot of casual acquaintances and one great friend, but "prefers being alone to being with a lot of people". Will be very sorry to leave school—"has got into a rut and hates to leave ruts". Wants to go to the University and get into the Civil Service; "is not keen on the work for itself, only chose it because it is a good job—good prospects, steady job, good pay—and one ought to get one after a good education"; "otherwise would have acted or been a doctor"; "if she fails will be the inevitable secretary".

A clear case of both superiority and dominance drives was that of T1 in Group (*a*); she was physically deformed, and compensated for this by an unusual degree of self-assurance, which might tend towards aggressiveness:

She likes any school subject "when she can do it, but not when she can't"; likes biology and chemistry—"wants to get on to dissecting—it would be lovely"; also likes history—"likes to hear what people have done". Is quite keen on getting a job if it is interesting—"something she really liked—not too monotonous"; wouldn't like to teach, she "would get so impatient she would throw things at them", but wouldn't mind "being a head mistress and running a school—would like organizing"; secretarial work might be "too routine—but nice to be a secretary to someone very interesting".

Display. The drive towards display, accompanied by the desire for social admiration, was not infrequently coupled with the superiority drive in individuals who not only wished to feel superior to others, but also wished to be looked up to and admired for it. Examples of this were B4 and B1 in Group (*c*) (p. 137) and B3 in Group (*b*) (p. 138). But there were also girls who cared little about excelling in itself so long as they could receive some recognition and applause from their contemporaries; in fact they were examples of vanity rather than of pride! Such cases were H4 in Group (*c*) and R1 in Group (*a*); they speak for themselves:

H4 likes dancing and gym. best, and had intended to take up ballet dancing professionally, had she been strong enough—"was entered for a ballet" and "used to do four hours (practising) a day". Is going to take up professional skating; "might try for a teaching job" or "exhibition skating"; "would rather do it casually, not as a permanent job"—"wouldn't like it so much if she had to do it". "Parents would like her to stay at home." At one time "thought of beauty culture—her father was quite willing to buy her a shop if she wanted one"; but "everyone seems to be doing it"—there are "too many shops"; she "sets people's hair at school—has the knack of it—it's great fun"; "would like doing the chemical side too, making beauty preparations, but it needs too much chemistry".

R1 likes debating—"likes arguing about something she knows about"; enjoys acting—did a lot at home before she came here to school—and dancing, especially tap-dancing—"but doesn't do it now"; likes playing games, but "hates netball—is not tall enough to catch the ball—gets knocked about—it's so senseless"; is captain of middle school tennis and hockey—was form representative last year—"had to tell the house captain how people play and arrange the team"—enjoyed it, but it was "awkward having to tell people they weren't to play". When she leaves would "like to dance for fun"—is "not good enough for a professional", which is "very hard work" too.

A rather similar case in Group (*b*) was that of L1, who desired admiration for her artistic skill and manual dexterity:

She likes needlework, especially embroidering "tea trays and table cloths". Wants either to take up hairdressing and be apprenticed to a shop in the town; or "would like to work in a florist's shop—likes arranging flowers".

Pleasure and variety. There was sometimes a close relationship between the drive towards display and the drive towards pleasure and variety. That is to say, the pleasure-seeker achieved the maximum of self-gratification with the minimum of effort in the society of congenial companions who approved and admired her behaviour and gratified her wishes. R1 in Group (*a*) and H4 in Group (*c*) were good examples of this link between pleasure seeking and display:

Thus R1 also likes biology—"it's much the nicest lesson because it's a new subject, not stale", they have only started on it this term. Her mother has suggested her taking up domestic science but she doesn't want to do this—"hates physics and chemistry"; hasn't thought much about jobs, and "gets bored because Mother is always talking about it"; would like to be a secretary if she could live in the country—but "also likes to be within reach of amusements".

H4 also thinks that "term is so monotonous—is rather fed up with it here—will be glad to leave". Prefers town to country in the holidays—"has two brothers and goes out with them a lot."

A rather similar case was that of M1 in Group (*a*):

She likes history and drawing, but "doesn't like any other lessons"; likes playing games and dancing, going to theatres and cinemas; prefers holidays to term because "there is not so much work to do"; is "rather keen on taking up dancing"—ball-room dancing, not stage dancing; would "like to go abroad, when she leaves, before starting a job"; parents want her to do domestic science—"would much prefer dancing"; would hate teaching domestic science.

In some other cases the pleasure drive appeared mainly as the antithesis of the activity drive. Thus the individual who was motivated by the former and not the latter was characterized less by a sustained effort to obtain pleasure and amusement than by a sort of restlessness and an inability to persevere in any activity which demanded thought, energy or concentration:

Thus in Group (*c*) W2 likes games and dancing best, but "likes everything except maths. and science"—"hasn't got a mind that way"; is specializing in domestic science—likes cooking and dressmaking best—"housewifery is harder work". Is a sub-prefect and "has to try to keep the others in order"—"keeping one's own friends in order is difficult". Prefers holidays to term—"gets fed up towards the end of term—it goes on too long"; likes outdoor life in the holidays "but not being too hearty"; it is dull being at the seaside in the spring "because no one else is there" and she's "not terribly good at golf". Will be "sorry in a way to leave"—it will be an "awfully strange feeling"; but will be "glad to get away from the routine"—"likes variety". Had thought of taking up physical training when she leaves; "might have done it if she'd had to do something", but it would be a "nuisance to go to a Physical Training College for two years" and she would "hate to teach it"—it would be "awful to be a school teacher". Is going to do a course of training in domestic science; "has thought of a career but there seems to be nothing for her to do but domestic science", though she "may possibly get bored at home doing nothing"; her "parents want her to stay at home"—"Mother wants her to take over the housekeeping", but she doesn't want to do this.

A1 (Group (*c*)) likes tennis and dressmaking; likes French and history—especially "learning new things". Is a sub-prefect—finds the duties "rather a nuisance—boring—doesn't like bossing others about". Prefers holidays to term—likes to "get away from the routine—hates coming back to school", it is a "grind after the holidays"; will be "sorry to leave in some ways" but "glad to get away from the discipline". "People at school want her" to go to College, but she doesn't want to; "would like to have something to do, not sit at home all day"; "might go in for dressmaking"; "would really like to live at home and work at it at her leisure—a job ties you rather".

S1 in Group (*b*) was almost completely sluggish in her inactivity; it was this that accounted for her wish to stop at home, not timidity, in which she was completely lacking:

She likes needlework—"making frocks for herself"—and cooking; out of school she likes reading novels; would rather be at home than at school—"gets tired of school"; when she leaves will "stop at home for a bit with mother"; may go out to work later, but will "wait and see"—is not keen to; will be "glad to go home" when she leaves.

VOCATIONAL PREFERENCES OF SCHOOL-LEAVERS IN A SCOTTISH INDUSTRIAL AREA[1]

By HELEN PALLISTER

From the Psychological Laboratory, St Andrews

I. *Introduction* (pp. 144–145).
II. *Purpose of the investigation* (pp. 145–146).
III. *Girls* (pp. 146–156).
  A. *First vocational preferences* (pp. 146–148).
  B. *First vocational preferences compared with industries actually entered* (pp. 148–149).
  C. *First vocational preference in relation to second vocational preference* (pp. 149–151).
  D. *First vocational preference in relation to education* (pp. 151–153).
  E. *First vocational preference in relation to club membership* (pp. 153–154).
  F. *First vocational preference in relation to interests* (pp. 154–156).
IV. *Boys* (pp. 156–163).
  A. *Vocational preferences* (pp. 156–158).
  B. *Vocational preference in relation to education* (pp. 159–160).
  C. *Vocational preference in relation to club membership* (p. 161).
  D. *Vocational preference in relation to interests* (pp. 161–163).
V. *Summary* (pp. 163–166).
  (1) *Problem and procedure* (p. 163).
  (2) *Results* (pp. 163–166).
VI. *Conclusion* (p. 166).
  References (p. 166).

I. INTRODUCTION

EVERY year thousands of school-leavers in Britain face the problem of entering the ranks of employed workers. The vocational psychologist is interested in knowing for what occupations these juveniles express preferences. The vocational preference, existing as a positive attitude toward a certain occupation, is one factor of a pattern in which function together the economic condition and industrial structure of the area, the economic and social status of the juvenile's parents, the mental level, special abilities, personality attributes and interests of the juvenile, his knowledge of, or misconceptions about various occupations, his conscious or unconscious acceptance of a certain role in society, and the attitude of others toward the juvenile. The analysis of the data presented in this paper is both of theoretical and practical importance. It is theoretically interesting to relate the factors of education, personality and interest to the vocational preferences expressed, and to discover whether the juveniles' pattern of preferences coincides with the occupational pattern in the area. It is also important for practice, as these children, and thousands like them, are not mere case numbers in the files of the Ministry of Labour. They are living human beings whose occupational fitness and adjustment are the concern of society and the special problem of vocational psychologists as applied scientists.

---

[1] This paper was read to the Psychology Section of the British Association for the Advancement of Science, Blackpool, 1936, as one in a series of *Co-ordinated Social Researches in a Scottish Area*. The first paper in this series, by O. A. Oeser(2), describes the general organization and methods of the researches, which are financed by special grants from the Pilgrim Trust and the Medical Research Council.

II. PURPOSE OF THE INVESTIGATION

This paper is concerned with a statistical analysis of the vocational preferences of 925 juveniles in Dundee. These juveniles, 478 girls and 447 boys, left the central schools of the area in June 1936. The data on which the analysis is based come directly from form E.D. 210 X of the Ministry of Labour. This form consists of a small, four-page leaflet divided into two sections: pp. 1 and 2, filled in by the prospective school-leaver a few weeks before he leaves school; pp. 3 and 4 filled in by the teacher at the same time. The first section supplies, among other information, that concerning:

(*a*) first vocational preference of the child;
(*b*) second vocational preference;
(*c*) intention of attending evening continuation classes;
(*d*) club membership;
(*e*) interests and hobbies.

The second section will not be considered in any detail in this paper. The only information from it which has been analysed is that concerning the educational status of the child. The educational status of the child as given on form E.D. 210 X was later checked against the official list of school certificates awarded, and any necessary corrections were made, since it was not possible in all cases for the teacher to predict correctly whether or not a child would obtain a Day School Lower or Day School Higher Certificate.[1]

The sexes have been treated separately. This is essential; as the environment does not always press uniformly on the two sexes. It is particularly important for the vocational psychologist to determine the effect of differential pressure on the juveniles, and also to show how, when common environmental factors exist without differential pressure, the sexes resemble each other in status or behaviour.

III. GIRLS

A. *First vocational preferences*

Table I shows the first vocational preferences of 478 girls. Vocational preference here means the occupation or job which the juvenile writes on the form as the occupation he or she would like to enter.[2] The concept of vocational preference implies a vocational wish restricted by the individual's conscious or unconscious awareness of environmental or even hereditary factors which circumscribe the range of possible vocational choice. A girl might wish to be an engineer, but never expresses a vocational preference for this occupation, because of the social pattern of this area in which boys may be engineers, but girls are not engineers. For both boys and girls, vocational wishes may be restricted by such factors as the economic conditions of the parents, opportunities for training and social attitudes toward different occupations.

The vocational preferences of girls have been classified in ten occupational categories, shown in Table I:

(1) mill (mainly jute spinning, but may include spinning of flax or hemp);

(2) factory (jute, and possibly flax weaving);

(3) domestic;

(4) laundry (manual work in commercial laundries);

(5) preserves (dealing with the making of sweets or the canning of fruits or vegetables);

(6) printing (e.g. hand colouring cards, tying ribbons, assembling calendars and similar work in printing establishments);

---

[1] The author gratefully acknowledges the permission granted by the officials of the Ministry of Labour to use their files to obtain the data for this investigation. She also welcomes this opportunity of expressing thanks to her colleagues, E. L. Trist, H. Hillmann, and W. O'D. Pierce, for very considerable aid in transcribing the raw data.

[2] Question 8 on the form reads: "What kind of work would you most like to have? If you have any *special* reasons for your choice, state them."

(7) dressmaking;
(8) shop assistant (including message girl);
(9) office;
(10) miscellaneous occupations.

Some of these categories contain very few individuals, but the categories have nevertheless been included in the classification, because these occupations are entered by girls in this area.

Table I. *Girls. First vocational preferences*

Vocational preference	Frequency of preference	% of total
Mill	21	4·4
FACTORY	**128**	**26·8**
Domestic	18	3·8
Laundry	10	2·1
Preserves	5	1·0
Printing	6	1·3
Dressmaking	15	3·1
SHOP ASSISTANT	**145**	**30·3**
OFFICE	**94**	**19·7**
Miscellaneous	36	7·5
Total	478	100·0

Table I shows that the girls have three dominant vocational preferences, which are, in order of decreasing popularity: shop assistant (30·3 %), factory work (26·8 %) and office work (19·7 %). These three occupations together account for 76·8 % of all vocational preferences. It is the writer's opinion that the most popular occupation, namely, shop assistant, appeals to the juvenile because of supposed personal or social advantages not offered by any of the other occupations. The shop assistant can wear more attractive clothes and keep herself neater than the girl in any of the other occupations except the office worker. She has a greater advantage than the other girls, including the office girls, in meeting people, and for small shops high educational attainments are not required as they are for office workers. Until she has had some experience in the work, the juvenile does not seem to see the disadvantages of shop work, such as the long hours, and the possibility of being permanently laid off at the age of 18.

Factory work, the second choice in popularity, also carries with it a prestige value. In this area spinners and weavers now work under conditions that are objectively similar. Yet weaving is considered superior to spinning, although to an outsider both kinds of work appear equally dirty and unpleasant. The prestige value of weaving continues to exist, although one of the original factors at the root of this prestige, better wages for the factory worker, no longer exists. Hence we find that only

4·4 % prefer a mill job, whereas 26·8 % express a preference for factory work.

B. *First vocational preferences compared with industries actually entered*

The vocational psychologist should know not only the popularity of occupations in the area, but also the distribution of workers in these occupations. It has not been possible in all cases to compare the frequency of vocational preferences directly with the absorptive capacity of the different occupations, because the official figures of the Ministry of Labour report the numbers employed only on an industrial basis. By adding together the number preferring mill and factory work a category is obtained comparable to the industrial categories of jute, flax and hemp. The category of vocational preference known as shop assistant is comparable to the industrial category of distributive trades. Preserves, printing and laundry categories can also be compared with similar industrial categories. Figures for industrial categories comparable to the domestic, dressmaking, office and miscellaneous categories of vocational preference cannot be obtained.

Table II. *Girls.* Comparison of industrial employment,* 1936, *with first vocational preference*

	Employed workers				First vocational preference	
	Under 16 years of age $N=2749$		16–18 years of age $N=2737$		14–16 years of age $N=478$	
Industry	No.	%	No.	%	No.	%
Jute, flax and hemp industries	1175	42·7	1404	51·3	149	31·2
Distributive trades	446	16·2	565	20·6	145	30·3
Confectionery and other food industries	381	13·8	97	3·5	5	1·0
Printing	209	7·6	96	3·5	6	1·3
Laundry service	52	1·9	42	1·5	10	2·1
Total	2263	82·2	2204	80·4	315	65 9
Other industries	486	17·8	533	19·6	163	34·1
Grand total	2749	100·0	2737	100·0	478	100·0

* The raw figures for employed workers are taken from the official figures of the Ministry of Labour in the area.

Table II shows a comparison of available figures. Important facts brought to light are:

(1) A smaller proportion of girls prefer to enter the jute, flax and hemp industries than the proportion of juveniles employed in these industries. This difference is statistically significant[1].

---

[1] A difference is considered significant, in this paper, only if the difference between the two percentages divided by the sigma of the difference is 3 or greater than 3.

(2) A larger proportion of girls prefer to enter the distributive trades than the proportion of juveniles employed in these trades. This difference is also significant.

(3) Almost twice as great a proportion of girls prefer to enter 'other industries' as the proportion of workers in 'other industries'.

The general conclusion is that girls' vocational preferences do not vary as the absorptive capacity of the industries in the area. Occupational maladjustment may arise from this situation, since many workers are being forced into the kind of work they do not wish to enter.

Although those employed as shop assistants in this area do not constitute the largest numerical group, the distributive trade is one of the most important trades for juveniles in the country as a whole. Gollan (1) states that "distribution has become the greatest single industry for youth". This statement is based upon figures from the *Ministry of Labour Gazette* for July 1935. For Scotland as a whole, 12 % of juveniles were distributive workers at that date. For the girls of the nation as a whole Gollan finds the following industries in order of decreasing number of juveniles employed: distribution, textiles, clothing, food, drink, tobacco, other metal, paper and printing. In Dundee, textiles come first and distribution second in order of the number of female juveniles employed. Together they absorb 58·9 % of female workers under 16 years of age, and 71·9 % of female workers between 16 and 18 years of age.

C. *First vocational preference in relation to second vocational preference*

Instead of an analysis showing all the second vocational preferences in relation to first preferences, a much simpler analysis has been made. Table III shows for the different occupational groups based on first vocational preference the number and percentage who (*a*) give no alternative preference, (*b*) write 'anything', and (*c*) give a second specific vocational preference.

For the group as a whole nearly half of the children, 49·2 %, fail to mention a second specific vocational preference. It is for this reason that no attempt has been made to analyse the classified second vocational preferences in relation to first preferences.

In the column headed 'anything' all first preference groups show a percentage exceeding 16 except the office group. This figure, 3·2 %, for the office group points to the definite intentions of this group which has in school received specific vocational training in commercial courses.

Table III. *Girls. First vocational preference in relation to second preference*

First vocational preference	No alternative		'Anything'		Specific second preference		Total	
	No.	%	No.	%	No.	%	No.	%
Mill	4	19·0	8	38·1	9	42·9	21	100·0
Factory	13	10·1	57	44·6	58	45·3	128	
Domestic	2	11·1	8	44·4	8	44·5	18	
Laundry	1	10·0	6	60·0	3	30·0	10	
Preserves	1	20·0	2	40·0	2	40·0	5	
Printing	1	16·6	1	16·6	4	66·8	6	
Dressmaking	3	20·0	4	26·7	8	53·3	15	
Shop assistant	33	22·7	36	24·8	76	52·5	145	
Office	32	34·0	3	3·2	59	62·8	94	
Miscellaneous	8	22·2	12	33·3	16	44·5	36	
Total	98	20 5	137	28·7	243	50·8	478	100·0

There is a significant difference between the percentage of girls giving office work as first choice who state they will take 'anything' as second choice and the percentage of girls naming mill, factory, domestic, laundry, and shop assistant as first choice who state they will take 'anything'. There is also a significant difference between the factory and shop assistant group, the factory group being more willing than the shop assistant group to take 'anything'.

More than one-third of the mill group, more than two-fifths of the factory and domestic groups, and approximately one-quarter of the dressmaking and shop assistant groups are willing, so they state, to try 'anything' if they are unable to obtain their first choice of work. These figures show that for the mill, factory, domestic, laundry and preserves groups less than half the girls name a second specific vocational choice. At least two factors produce this response:

(1) The children in these groups leave school, on the average, at an earlier age than those in the remaining occupational groups. They are therefore not so likely to know the advantages and disadvantages of occupations, or the requirements of different occupations, and hence give no alternative choice or say that 'anything' will suit them.

(2) The economic conditions of their homes make it imperative that they get a job immediately, regardless of the nature of the work. It should be noticed that the percentage of those who say 'anything' for the mill, factory, domestic, laundry and preserves groups is much larger than the percentage of 'no alternatives' for the same groups.

Of the three most popular occupational groups, judged by first choice, office has the largest proportion naming a second specific vocational choice, next comes shop assistant, and finally factory. The differences in percentages are not, however, completely reliable statistically.

Of the total number of girls who do name a second vocational choice, 105 (22 %) prefer the occupation of shop assistant. Shop assistant is thus the most popular second as well as first choice.

D. *First vocational preference in relation to education*

(1) *Educational status.* By educational status is meant one of the following. The child:

(*a*) has been in a class for retarded pupils because of physical or mental defects—Retarded group;

(*b*) is supposedly normal, but is leaving school without a certificate—Normal group;

(*c*) has obtained a Day School Lower Certificate but no higher certificate—D.S.L. group;

(*d*) has obtained a Day School Higher Certificate—D.S.H. group.

[D.S.L. certificates are awarded for satisfactory scholastic attainment in central schools, usually at the age of 14. D.S.H. certificates are awarded as the result of special examinations and are obtainable one year after the D.S.L.]

Of the total group the children who obtain a D.S.L. but no higher certificate comprise 27·8 % (see Table IV). The D.S.H. category comprises 16·1 %. There are 37·9 % in the Normal and 18·2 % in the Retarded category. If the numbers in the D.S.H. and D.S.L. categories are combined to find the total percentage of children who leave school with at least a D.S.L. Certificate the figure is found to be 43·9 %. All holders of a D.S.H. must have previously obtained a D.S.L.

Table IV. *Girls. First vocational preference in relation to education*

Vocational preference	Educational status										Intention of attending continuation classes	
	Retarded		Normal		D.S.L.		D.S.H.		Total			
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Mill	11	52·4	8	38·1	2	9·5	0	0·0	21	100·0	6	28·6
Factory	35	27·3	54	42·2	39	30·5	0	0·0	128		28	21·9
Domestic	9	50·0	5	27·8	4	22·2	0	0·0	18		6	33·3
Laundry	4	40·0	5	50·0	1	10·0	0	0·0	10		3	30 0
Preserves	1	20·0	2	40·0	2	40·0	0	0·0	5		1	20·0
Printing	1	16·7	3	50·0	2	33·3	0	0·0	6		4	66·6
Dressmaking	2	13·3	10	66·7	2	13·3	1	6·7	15		6	40·0
Shop assistant	12	8·3	63	43·5	60	41·3	10	6·9	145		76	52·4
Office	0	0·0	17	18·1	14	14·9	63	67·0	94		91	96 8
Miscellaneous	12	33·3	14	38·9	7	19·5	3	8 3	36		13	66·1
Total	87	18·2	181	37·9	133	27·8	77	16·1	478	100·0	234	48·9

Consider the four educational categories separately, again. Several points should be noted. Not until the dressmaking group is reached does anyone appear with a D.S.H. Of the shop assistant group a small percentage have a D.S.H. Of the office group 67·0 % have a D.S.H. These 63 girls constitute 81·8 % of the 77 girls who obtain a D.S.H. They are intellectually the *élite* of this school population. However, it should be remembered that a number of children are obliged to leave school before they have had the necessary training to take the examination.

In all the occupational groups some children have a D.S.L. Mill (9·5 %) and laundry (10 %) have few D.S.L. pupils, whereas factory (30·5 %), printing (33·3 %), shop assistant (48·2 %) (41·3 + 6·9) and office (81·9 %) (67 + 14·9) embrace increasingly larger numbers of pupils with a D.S.L. at least. The office significantly exceeds all other groups. In all occupational groups, with the exception of the office group, from 27·8 to 66·7 % of the pupils are Normals with no certificate. Of the office group only 18·1 % have no certificate. With regard to the Retarded category, of the mill group 52·4 % are retarded, whereas no one in the office group is placed in this educational category. Spinning is not only unpopular as an occupation, but more than half of those who express a preference for it are retarded because of either physical or mental defects. Of the domestic group, half are also retarded. The printing, dressmaking and shop assistant groups show progressively small percentages of retarded pupils in them. The shop assistant and office groups differ significantly from the mill, factory and domestic groups in having a smaller percentage of retarded pupils in them.

A relationship thus appears between first vocational preference and educational status. It appears most markedly if the highest and lowest categories of educational achievement are considered in relation to vocational choice. Those who express a preference for occupations of a high prestige value have higher educational attainment, as a group, than those choosing a job lower in social prestige.

(2) *Continuation classes.* Table IV also shows the relation of first vocational preference to intention of attending continuation classes. The average for the total group of 478 girls is 48·9 %. This figure is almost double that given by Gollan (1) as the percentage of school-leavers who actually begin evening continuation classes. Gollan sets the percentage of entrants to continuation classes at 25.

The various occupational preference groups show a marked difference in the percentage of those who state that they intend to attend such classes. One third or less of the mill, factory, domestic and preserves groups plan to attend continuation classes. The dressmaking, shop assistant, miscellaneous, printing and office groups, in order of increasing percentages, plan to attend these classes. The office group stands out again as intellectually superior. Not only is it characterized, on the one hand by no retarded members, and on the other hand by the greatest number with D.S.H. certificates, but also by the greatest number planning to continue their education. In part, at least, economic factors explain these differences. Girls leaving at 14 years of age who require an immediate job, realistically pick one that does not make any great demands upon ambition, a good educational status or further training. Those leaving with a better education know that financial promotion comes more easily to those who continue to excel in educational competition with others. Thus girls planning to enter shops learn that the better shops, where at least they will secure greater prestige, demand girls with a D.S.H. Those planning to enter clerical posts know that they will fare better if they take the Chamber of Commerce Examination. It is also true that the girls who remain at school to get a D.S.H. are more interested in school work than pupils who voluntarily leave school at 14, even when a job is not urgently required. Hence these better educated groups plan to continue their education further in evening classes.

E. *First vocational preference in relation to club membership*

The figures for club membership in relation to first vocational preference are as follows: Of the total group, 42 % are club members. The domestic and printing groups have the lowest percentage of club members of any of the occupational groups—only 16·7 % each. The mill group is also low with only 23·8 % who are members of clubs. The shop assistant and office groups exceed the average of the total group in membership (43·3 % and 56·4 % respectively). The only differences in percentages of club members which are significant are those between the mill and the office groups, and between the domestic and the office groups. The most striking feature about these figures is not the relation of club membership to type of job preferred, but the small number of club members in each of the occupational preference groups. This phenomenon may be explained by one or both of the following factors:

(*a*) Economic. Clubs cost money.

(*b*) Cultural. In this area socially organized and integrated groups are not much encouraged within the school system. For example, in one large central school, the only existing clubs are football clubs for boys, the orchestra and the choir. No athletic clubs exist for girls. There are no literary, dramatic, debating, science or nature study clubs such as are common in schools elsewhere. It is felt by the headmaster of this school that the pupils in central schools are too young for an extensive system of clubs to be developed.

That club activities on school premises are not frequent is also shown by the fact that the gates of the school grounds are locked soon after four o'clock, the time when day classes at the school end.

F. *First vocational preference in relation to interests*

Table V shows the figures for nine classified interests in relation to vocational preference. For each individual in each of the various occupational preference groups only the number of different kinds of interests mentioned have been recorded. Thus, although a girl may have mentioned three different sports which she likes, but only one kind of sewing, her interest in athletics is recorded only once as is her interest for sewing. The nine classified interests are: athletics, music, sewing, etc. (including embroidery, knitting and crochetting), domestic activities (such as cooking and housekeeping), art, reading, collecting, miscellaneous interests, and no interests.

The figures at the bottom of the table show that there are three dominant kinds of interests among girls: athletics, sewing, etc., and reading, in decreasing order of importance. All the other interests are very infrequent.

Now compare the occupational preference groups by interests. First, the average number of interests, given in the column headed *A*/*B* is low for the laundry, miscellaneous and factory groups. It is the highest for the printing and office groups, but not even for these groups does the average number of interests reach three. [This is to be expected in view of the fact that only three dominant interests are to be found in the total group of girls.] A few individuals in all groups, except the laundry and dressmaking groups, both small groups, fail to mention any interests, although it is psychologically improbable that these individuals are really lacking in interests outside their school work. It is more likely that their interests may be ones which they are not so ready to admit while in school, interests such as the cinema. All the recorded interests of the girls are such that they would meet with the approbation of the school authorities or are vocationally desirable. Thus it seems that the school environment, and the fact that form E.D. 210 X, after being filled in, is submitted to the school authorities, directly influence the enumeration of interests. Here we see the potent effect of one institution in conditioning the interests and/or the expression of interests of those coming under its control.

Table V. *Girls. First vocational preference in relation to interests*

Vocational preference	Athletics		Music		Sewing		Domestic activities		Art		Reading		Collecting		Miscellaneous activities		None		Total no. of interests A	Total in occupational group B	A/B
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	A	B	A/B
Mill	8	38·1	0	0·0	11	52·4	5	23·8	1	4·8	3	14·3	0	0·0	1	4·8	1	4·8	30	21	1·43
Factory	61	47·7	2	1·7	44	34·4	15	11·7	3	2·3	30	23·4	2	1·7	3	2·3	12	9·4	172	128	1·34
Domestic	8	44·4	0	0·0	8	44·4	3	16·7	0	0·0	8	44·4	0	0·0	1	5·6	2	11·1	30	18	1·67
Laundry	6	60·0	0	0·0	2	20·0	0	0·0	0	0·0	4	40·0	0	0·0	1	10·0	0	0·0	13	10	1·30
Preserves	2	40·0	1	20·0	3	60·0	0	0·0	0	0·0	1	20·0	0	0·0	1	20·0	1	20·0	9	5	1·80
Printing	3	50·0	0	0·0	2	33·3	0	0·0	2	33·3	4	66·7	0	0·0	1	33·3	1	33·3	13	6	2·17
Dressmaking	8	53·3	3	20·0	8	53·3	0	0·0	0	0·0	5	33·3	2	13·3	0	0·0	0	0·0	26	15	1·73
Shop assistant	88	60·6	5	3·5	58	40·0	9	6·2	3	2·1	60	41·4	1	0·7	3	2·1	11	7·6	238	145	1·64
Office	63	67·0	29	30·9	39	41·5	0	0·0	1	1·1	62	66·0	12	12·8	8	8·5	5	5·3	219	94	2·33
Miscellaneous	15	41·7	1	2·8	13	36·1	2	5·6	2	5·6	6	16·7	0	0·0	5	13·9	3	8·3	47	36	1·31
Total	262	54·8	41	8·6	188	39·3	34	7·1	12	2·5	183	38·2	17	3·6	24	5·0	36	7·5	797	478	1·67

The percentages read horizontally show the percentage of individuals recording an interest in each occupational group. These interests are treated as independent variables. Therefore, since an individual may have more than one interest, the percentages' read horizontally do not total 100·0%.

The pattern of interests differs somewhat for the different occupational groups. All groups show the three dominant interests to some extent, but in varying proportions. However, only in reading does any reliable difference of interest exist between the occupational groups. The shop assistant group is more interested in reading than either the mill or factory groups, and the office group is more interested in reading than the mill, factory, dressmaking or shop assistant groups. These differences are in accord with expectation. The mill group, which has the lowest percentage of girls interested in reading, has the highest proportion of retarded pupils in it, whereas the office group, and to a less extent the shop assistant group, have a high proportion of well-educated individuals in them. Also this interest is indirectly related to occupational adjustment. The necessity for dealing with verbal symbols is unimportant for a girl intending to enter a mill or a factory, whereas it is an essential part of the daily routine of an office girl, and is also necessary for the shop assistant.

Some of the recorded interests seem not to accord so well as the above examples with the vocational preference expressed by the child. Of those in the domestic group only 16·7 % mention such interests as cooking and housekeeping, which are classified under the head of domestic interests in the table. This, coupled with the fact that a large proportion, 50 %, of the girls in this group are retarded, points to an unwise choice of occupation. It is rather surprising, too, that only 53·3 % of the dressmaking group mention an interest in sewing. The fact that 30·9 % of the office group mention music as an interest is quite striking. Actually, these twenty-nine girls in the office group constitute 70·7 % of the group of 41 girls who list music among their interests. Music as a cultural interest is probably related here to the superior socio-economic status of homes from which girls in the office group come. These are girls whose parents can afford to keep them in school until they obtain a fair commercial training. We may suppose, therefore, that money for such cultural advantages as music is more readily available in the homes of these girls. The monetary factor cannot, of course, be the whole explanation, but it is still true that the interests of those with aptitude for music is strengthened through musical education.

IV. BOYS

A. *Vocational preferences*

The same analyses are made for boys as for girls, except that the relation of industrial employment to vocational preferences, and the relation of first to second vocational preference, have not been analysed.

Table VI presents an analysis of the classified vocational preferences of boys. Whereas the girls mention only twenty-six jobs which are readily classified into ten occupational categories, the boys mention seventy-eight jobs. This restriction of the vocational preferences of the girls is a very interesting phenomenon. Environmental pressure is continuously bearing on females, forcing them into narrow vocational channels. Women who wish to enter 'men's jobs' are frowned upon, refused for such jobs, or admitted grudgingly at wages lower than those paid to men. The effect of this pressure has been so marked even on the girl no older than 14 or 16 years of age that she responds positively, as indicated by the preferences expressed, to only one-third of the number of jobs to which a boy responds.

Table VI. *Boys. Vocational preferences*

Vocational preference	Frequency of preference	% of total
Unskilled	74	16·6
Semi-skilled	35	7·8
Total	109	24·4
Skilled trade:		
Distributive	40	8·9
Transport	26	5·8
Building	79	17·7
Furnishing	4	0·9
Engineering	84	18·8
Miscellaneous	28	6·3
Total	261	58·4
Business	40	8·9
Professional	12	2 7
Total	52	11·6
Miscellaneous unclassifiable	25	5·6
Grand total	447	100·0

Because of the large number of occupations which boys mention it has been necessary to group these occupations differently from the classification for girls. Boys' preferences have therefore been classified as: unskilled, semi-skilled, skilled, under which are six subheadings; then business, clerical posts, professional (which also includes semi-professional occupations), and miscellaneous unclassifiable occupations. The most popular class of occupation is the skilled trade group. Of all the boys, 58·4 % mention one or another of the skilled trades. In this area only approximately 10 % of boy workers (250 out of 2547) actually entered an apprenticeship in 1936. A significant difference exists therefore between the number who wish to enter a skilled trade and the number of those who are eventually accepted for one.

Of the skilled trades, engineering ranks first in popularity, with building a close second. The next highest percentage occurring in the table is 16·6 %, the figure for those mentioning some occupation of an unskilled kind (such as railway cleaner, sawmill, lodge boy, etc.). Some of these boys intend to take up other work later, but as no other preference has been stated the only job mentioned as the boy's preference has had to be accepted. Ambiguity arises here in the interpretation of the vocational preferences written on the form. Boy's first and second vocational preferences cannot be clearly distinguished as they are for girls. Therefore a table showing the percentage of second vocational preferences in relation to first preference is not given. It is the apprenticeship system with the pre-apprenticeship years, 14–16 years of age, taken up in temporary work, that renders any clear distinction between first and second preferences for boys impossible. Whenever a boy mentions an apprenticeship as his goal at the age of 16, this is classified as his vocational preference, although on the line calling for first vocational preference he may have written the name of a temporary job, such as errand boy.

Further differences between the kinds of jobs chosen by boys and girls are now apparent in the nature and length of training required to fit a person for the job. Of the boys, 58·4 % wish to enter an apprenticeship. While it should be noted that the length of time spent in an apprenticeship to-day bears no direct relation to the number or complexity of the skills acquired during the training period, nevertheless the trained apprentice does acquire numerous skills not possessed by him at the time he leaves school. Even more important is the fact that he has prospects of advancement. In the case of all the occupations which the central school girls choose to enter, the training period is relatively short. Furthermore it is not intended that a girl should acquire such technical knowledge of the job at which she works as a boy would acquire. For example, no matter how skilled a girl weaver may become, she is not expected to understand the treatment of jute for various industrial purposes, nor is she required to master the theory or practical mechanics of her loom. Therefore, although a boy may advance from apprentice tenter (mill mechanic) to manager of a works, the girl remains a weaver all her life.[1]

---

[1] Because of the greater diversity of occupations which boys as compared with girls enter, and the differences in skill required in various occupations within the same industry, it has not been possible to compare industrial employment of boys with their vocational preferences. Figures for industrial employment give no clue to whether the boy is doing unskilled or semi-skilled work, or whether he is a skilled apprentice or clerical worker.

B. *Vocational preference in relation to education*

Table VII shows boys' vocational preferences in relation to educational status and intention of attending continuation classes.

Table VII. *Boys. Vocational preference in relation to education*

	Educational status										Intention of attending continuation classes	
	Retarded		Normal		D.S.L.		D.S.H.		Total			
Vocational preference	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Unskilled	22	29·7	41	55·4	9	12·2	2	2·7	74	100·0	18	24·3
Semi-skilled	12	34·3	21	60·0	0	0·0	2	5·7	35		4	11·4
Skilled trade:												
Distributive	2	5·0	22	55·0	12	30·0	4	10·0	40		27	67·5
Transport	0	0·0	15	57·7	7	26·9	4	15·4	26		16	61·5
Building	17	21·5	32	40·5	25	31·6	5	6·4	79		39	49·4
Furnishing	0	0·0	3	75·0	1	25·0	0	0·0	4		2	50·0
Engineering	8	9·5	45	53·6	21	25·0	10	11·9	84		59	70·2
Miscellaneous	4	14·2	9	32·2	8	28·6	7	25·0	28		16	57·1
Total	31	11·9	126	48·2	74	28·4	30	11·5	261		159	60·9
Business	0	0·0	3	7·5	3	7·5	34	85·0	40	100·0	38	95·0
Professional	1	8·3	0	0·0	1	8·3	10	83·4	12		10	83·3
Miscellaneous unclassifiable	7	28·1	18	71·9	0	0·0	0	0·0	25		7	28·0
Grand total	73	16·3	209	46·7	87	19·5	78	17·5	447	100·0	236	52·8

(1) *Educational status.* The categories of educational status are the same as for girls. There are no significant differences between the total number of boys and girls in the four educational categories: Retarded, Normal, D.S.L., D.S.H.

It is very important to notice that of the total skilled trade group only 11·5 % obtain the D.S.H. A large percentage of boys, therefore, not only expresses a preference for occupations in which there are few openings, but also wishes to enter these occupations without adequate education. The question of evening continuation classes to bring these boys up to the necessary educational standard will be considered presently. Within the skilled trade group 25 % is the highest percentage of those having a D.S.H. A few in the unskilled and semi-skilled groups have a D.S.H. These figures, 2·7 and 5·7 % respectively, probably indicate not so much lack of ambition on the part of these boys as failure to state the life occupation they prefer. The highest percentage of boys holding the D.S.H. occurs in the professional group, 85, and in the business group, 83·4. If we add to those figures those for boys in the D.S.L. category we find that of boys planning to enter business 92·5 % have at least a D.S.L., and of the professional group, 91·7 % have obtained this certificate at least. Within the skilled trade group only approximately 40 % of the boys have at least a D.S.L. For some trades the percentage is lower.

Of the group of normal boys with no certificate the largest percentage occurs in the semi-skilled group. Even in the skilled trade group, however, from 32·2 to 75 % normals are without certificates. These percentages are far too high for boys hoping to enter a skilled trade. Of those wanting to enter business only 7·5 % are normals without a certificate; of the professional category no one is a normal without a certificate, but one boy is a retarded pupil. This retarded boy, regardless of the causal factor in his retardation, has no chance of entering work of a professional nature. The percentages in the skilled trade group of boys who are retarded again emphasize the failure of boys in this group to attain the required educational standards.

It is clear that the majority of these school leavers do not express a preference for occupations for which they are educationally prepared. There is a smaller discrepancy between the educational status and the educational requirements of the occupations preferred by the girls than there is for the boys, because the occupations preferred by the girls do not, on the whole, demand high educational qualifications.

(2) *Continuation classes.* Since the majority of boys need further education to fit them for the occupations they want to enter, examine now the figures for those intending to attend continuation classes. Even if all the boys in the skilled trade group who state that they intend to attend continuation classes actually do so, those who have not got certificates when they leave school will not all obtain them at the continuation school. 88·5 % of the skilled trade boys have not got a D.S.H. Only 60·9 % of the group plan to attend continuation classes. Some of these will be undoubtedly boys who already have the D.S.H., and many boys who attend the classes will fall out during the year.

The two vocational preference groups having the highest percentage of boys with D.S.H. certificates are also the ones with the highest percentage intending to attend continuation classes. Of the business group, 95 % intend to attend such classes, and of the professional group 83·3 %. The engineering group with 70·2 % ranks next. Of the total number of boys, 52·8 % state that they plan to attend continuation classes. This figure slightly but not significantly exceeds the percentage for girls, 48·9.

C. *Vocational preference in relation to club membership*

With regard to the club membership of boys, the figure for all the boys, 40·1 %, does not differ significantly from that for girls, 42·0 %. There is no vocational preference group without some club members in it. This is as true for boys as it is for girls. The smallest percentage is found in the unskilled group, 25·7. The largest percentage is found in the professional group, 58·3. Of those in the business group, 52·5 % are club members. The differences between these groups are however not significant. There is not therefore any clear relation between the occupational level and club membership nor between club membership and the nature of the job, i.e. whether the job involves dealing mainly with people or with things.

D. *Vocational preference in relation to interests*

Table VIII shows boys' interests classified in relation to vocational preference. The classification of interests is similar for boys and girls. The domestic and sewing interests of girls do not appear in the boys' table, but instead there are woodwork and engineering interests. For boys the chief interest is athletics. The cycling interest belongs in the same category as the other sports. It is put into a category by itself, because in addition to being a sport, it serves the function of equipment for a beginner's job in the distributive trades. It is interesting to notice, in this connexion, that less than one-third of the boys who mention the distributive trade as their vocational preference have an interest in cycling. Of course, a boy may be able to cycle, and therefore obtain a job as a message boy, without having any great interest in cycling as an activity.

It is interesting to compare the interests of boys and girls. For girls, besides athletics, there are two other interests of considerable popularity, i.e. sewing and reading. Aside from athletics, other interests of boys are infrequently reported. Only 11·4 % of boys express an interest in reading. For girls the percentage is 38·2, and reading is the third most popular interest. This difference between the percentage of boys and girls interested in reading is statistically reliable. Among boys the professional group is the most interested in reading, but the percentage for this group, 16·7, exceeds only slightly that for the mill group of girls, 14·3, and the mill group shows less interest in reading than any of the other groups of girls. The fact that girls mature before boys may be one explanation of girls' greater interest in reading. Also the adult emphasis

Table VIII. *Boys. Vocational preference in relation to interests*

Vocational preference	Cycling		Athletics		Wood-work		Engin-eering		Music		Art		Reading		Col-lecting		Miscel-laneous		None		Total no. of in-terests *A*	Total in occu-pational group *B*	*A/B*
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%			
Unskilled	28	37·8	42	56·8	3	4·1	1	1·4	0	0·0	2	2·7	5	6·8	3	4·1	4	5·4	12	16·2	100	74	1·35
Semi-skilled	10	28·6	22	62·9	2	5·7	1	2·9	0	0·0	1	2·9	4	11·4	2	5·7	1	2·9	8	22·9	51	35	1·46
Skilled trade:																							
Distributive	13	32·5	22	55·0	5	12·5	1	2·5	5	12·5	3	7·5	5	12·5	5	12·5	1	2·5	1	2·5	61	40	1·53
Transport	10	38·5	13	50·0	3	11·5	0	0·0	2	7·7	0	0·0	4	15·4	1	3·9	2	7·7	6	23·1	41	26	1·58
Building	27	34·2	36	45·6	12	15·2	1	1·3	1	1·3	1	1·3	9	11·4	4	5·1	3	3·8	11	13·9	105	79	1·33
Furnishing	2	50·0	3	75·0	0	0·0	0	0·0	1	25·0	0	0·0	0	0·0	0	0·0	0	0·0	1	25·0	7	4	1·75
Engineering	37	44·1	48	57·1	11	13·1	8	9·5	2	2·4	3	3·6	11	13·1	6	7·1	11	13·1	9	10·7	146	84	1·74
Miscellaneous	11	39·3	25	89·3	4	14·3	1	3·6	1	3·6	1	3·6	3	10·7	3	10·7	4	14·3	1	3·6	54	28	1·93
Total	100	38·3	147	56·3	35	13·4	11	4·2	12	4·6	8	0·3	32	12·3	19	7·3	21	8·0	29	11·1	414	261	1·59
Business	6	15·0	28	70·0	2	5·0	2	5·0	6	15·0	3	7·5	6	15·0	8	20·0	8	20·0	3	7·5	72	40	1·80
Professional	2	16·7	3	25·0	4	33·3	1	8·3	2	16·7	2	16·7	2	16·7	4	33·3	1	8·3	0	0·0	21	12	1·75
Miscellaneous unclassifiable	1	4·0	12	48·0	0	0·0	1	4·0	0	0·0	0	0·0	2	8·0	1	4·0	1	4·0	5	20·0	23	25	0·96
Grand Total	147	32·9	254	56·8	46	10·3	17	3·8	20	4·5	16	3·6	51	11·4	37	8·3	36	8·0	57	12·8	681	447	1·53

There are no significant differences between the occupational groups in the percentage of juveniles expressing an interest in athletics or any other activity.

See also footnote Table V.

on sports for boys may explain in part their concentration on this one type of interest.

Although only a small percentage of boys is interested in collecting, nevertheless the percentage for this interest, 8·3, is significantly greater than the percentage for girls, 3·6.

For boys the greatest average number of interests for any occupational group given in the column, *A/B*, is 1·93 for the miscellaneous skilled trade group. The business and professional groups are next highest, 1·80 and 1·75 respectively. None of the boys' vocational groups reaches or exceeds the figure, 2·33, for the office group of girls. This again shows the greater concentration of interest among the boys and the greater diversity among girls.

V. SUMMARY

(1) *Problem and procedure*

The vocational preferences of 925 juveniles, 478 girls and 447 boys, as recorded on form E.D. 210 X of the Ministry of Labour, are classified in categories based on occupational preference. These juveniles left the central schools of Dundee in June 1936. The sexes are treated separately. Raw figures and percentages are shown for each category of vocational preference.

Groups based on the various categories of vocational preferences are studied in relation to educational status, intention of attending continuation classes, club membership, and classified interests. Raw figures and percentages of these variables are reported for each category of vocational preference. For girls, analyses of industrial employment in relation to vocational preference and of second vocational preferences in relation to first preferences are also reported. Where a group is referred to as mill, office, skilled trade, etc. this means that it is a group based on the first preference of the individuals in the group for mill work, office work, or a skilled trade.

(2) *Results*

A. *Vocational preferences.* (1) Boys name three times as many jobs as girls and different jobs. Therefore the classification of boys' vocational preferences differs from that of girls. The first vocational preferences of girls are classified in the following ten categories: Mill, Factory, Domestic, Laundry, Preserves, Printing, Dressmaking, Shop Assistant, Office and Miscellaneous. The vocational preferences of boys are classified as follows: (1) Unskilled; (2) Semi-skilled; (3) Skilled Trade: (*a*) Distributive, (*b*) Transport, (*c*) Building, (*d*) Furnishing, (*e*) Engineering, (*f*) Miscellaneous; (4) Business; (5) Professional; (6) Miscellaneous unclassifiable.

(2) The three most popular vocational preferences of girls are shop assistant (30·3 %), factory (26·8 %), and office (19·7 %), in order of decreasing popularity.

(3) The most popular class of occupation among boys is the skilled trade category (58·4 %). Within this category, engineering is the most popular and building a close second.

B. *Industrial employment in relation to vocational preference.* (1) The frequency of occupational choices does not vary as the absorptive capacity of the industries in the area. A smaller proportion of girls prefers to enter the jute, flax or hemp industries than the proportion of female juveniles employed in these industries. A larger proportion of girls prefers to enter the distributive trades than the proportion of female juveniles employed in these trades. Other discrepancies also exist.

(2) A similar comparison for boys cannot be made from the available industrial figures. It has been shown, however, that 58 % of boys prefer to enter skilled manual trades whereas in 1936 only 10 % of the number of employed juveniles entered such apprenticeships.

C. *First vocational preference in relation to second preference* (*girls*). (1) 20·5 % give no alternative vocational preference. 28·7 % write 'anything' will suit them if they do not get work in line with their first vocational preference. This gives a total of 49·2 % who fail to name a second specific vocational preference.

(2) There is a significant difference between the percentage of girls preferring office as first choice (3·2) who state as second choice that they will take 'anything' and the percentage who prefer the following as first choice: mill (38·1), factory (44·6), domestic (44·4), laundry (60·0), and shop assistant (24·8).

(3) The factory group also significantly exceeds the shop assistant group in the percentage of girls willing to take 'anything' as a second choice of work.

(4) Explanations for these group differences are given in terms of knowledge and economic factors.

D. *Educational status.* (1) The juveniles are classified in one of the following four categories of educational status: (1) Retarded, (2) Normal (without a certificate), (3) D.S.L. (pupils with a Day School Lower Certificate), and (4) D.S.H. (pupils with a Day School Higher Certificate).

(2) Discrepancies exist between the educational level attained by individuals preferring popular occupations and the educational require-

ments of these occupations. These discrepancies are more marked for boys than for girls.

(3) Occupational preference groups highest in social prestige (office for girls, business and professional for boys) are characterized by a higher percentage of well-educated individuals, as judged by presence of certificate holders and absence of retarded pupils, than are occupational groups lowest in social prestige (mill for girls, unskilled and semi-skilled for boys).

(4) There are no significant differences between the percentages of boys and girls who are Retarded (girls 18·2, boys 16·3), Normal (girls 37·9, boys 46·7), D.S.L. (girls 27·8, boys 19·5), and D.S.H. (girls 16·1, boys 17·5).

E. *Continuation classes.* (1) A significantly larger percentage of juveniles preferring occupations highest in social prestige intend to attend continuation classes than of juveniles preferring occupations lowest in social prestige.

(2) Boys preferring to enter skilled trades do not intend to attend continuation classes in sufficient numbers to bring the group as a whole up to the required educational level for entrance to apprenticeships.

(3) There is no significant difference between the percentage of boys (52·8) and girls (48·9) intending to attend continuation classes.

F. *Club membership.* (1) In general, no significant relationship exists between membership in a club and vocational preference. Girls in the office group are the only exception. There is a significantly greater proportion of club members in the office group (56·4 %) than in the mill (23·8 %) or domestic (16·7 %) groups.

(2) No significant difference exists between the percentage of boys (40·1) and girls (42·0) who are club members.

G. *Interests.* (1) Boys and girls in general mention the same types of interests. Boys, however, mention woodwork and engineering not mentioned by girls, and girls mention sewing and domestic work not mentioned by boys. The other categories of interests are: athletics, music, art, reading, collecting, miscellaneous and no interest.

(2) Girls have three major interests: athletics, sewing and reading. Boys have only one chief interest, athletics.

(3) Both boys' and girls' recorded interests are such as to meet with the approbation of school authorities.

(4) In general no significant differences exist between the interests of the different groups based on vocational preferences. One exception is the reading interest of girls. Girls in the shop assistant and office groups

are significantly more interested in reading than girls in the mill or factory groups. The office group also exceeds the dressmaking and shop assistant groups in interest recorded in reading. Interest in reading is vocationally important for work in offices and shops.

(5) In some groups recorded interests point to an unwise vocational preference. Only 16·7 % of the domestic group and 53·3 % of the dressmaking group mention an interest in domestic activities or sewing respectively. Among boys, only 9·5 % of the engineering group record an interest in engineering as a hobby; only 15 % of the business group and 16·7 % of the professional group record an interest in reading.

(6) Girls are significantly more interested in reading than are boys. Boys exceed girls in the interest expressed in collecting, but the percentages of both sexes who mention collecting are small.

VI. CONCLUSION

The findings reported in this paper throw light upon some important aspects of juvenile life which are essential to vocational psychology. Shortcomings in the data are, in all probability, apparent to the reader as well as to the author. For example, the classification of vocational preferences, while designated as an occupational classification, is, in fact, partly occupational, partly industrial and partly a classification based on the range of skills required. Admittedly, it is not a scientific classification. Similar criticisms can be brought against the classification of other variables discussed above.

It has been the aim of this paper to set forth as adequately as possible whatever facts, relationships and psychological implications emerge from the available data. It must not be supposed that a psychologist would gather the original data as the Ministry of Labour is at present gathering it. Although relevant to this paper, a full criticism of form E.D. 210 X cannot be attempted here.

REFERENCES

(1) GOLLAN, JOHN (1937). *Youth in British Industry*. London: Gollancz.

(2) OESER, O. A. (1937). "Methods and assumptions of field work in social psychology. *Brit. J. Psychol.* XXVII, 343–63.

(*Manuscript received* 25 *March* 1938)

THE ULTIMATE BASIS OF AESTHETIC EXPERIENCE[1]

By W. R. D. FAIRBAIRN

I. *Introduction* (pp. 167–168).
II. *Unwarranted assumptions in aesthetics* (pp. 168–169).
III. *The beautiful in art and in nature* (p. 169).
IV. *The work of art as a 'found object'* (pp. 169–170).
V. *The historical origin of art* (pp. 170–171).
VI. *The genesis of the work of art* (pp. 171–173).
  (*a*) *The discovery of the object* (p. 171).
  (*b*) *Wish fulfilment* (pp. 171–172).
  (*c*) *The need for disguise* (p. 173).
VII. *The nature of aesthetic experience* (pp. 173–174).
VIII. *The world of art as a world of 'found objects'* (pp. 174–175).
IX. *The failure of aesthetic experience* (pp. 175–178).
  (*a*) *Over-symbolization and under-symbolization* (pp. 175–177).
  (*b*) *Form without content and content without form* (p. 177).
  (*c*) *The optimum of symbolic significance* (pp. 177–178).
X. *The principle of restitution* (pp. 178–179).
  (*a*) *The work of art as a 'restored object'* (p. 178).
  (*b*) *The optimal synthesis* (pp. 178–179).
XI. *Aesthetic theory* (pp. 179–180).
References (p. 181).

I. INTRODUCTION

THE present paper may be regarded as the sequel to a paper entitled "Prolegomena to a Psychology of Art", which I contributed to this *Journal* (1). There I pointed out that, whilst no psychology of art could be regarded as complete unless it were based upon the study of both the artist and the beholder,[2] nevertheless artistic activity can occur, theoretically at least, in the absence of any beholder, and that, therefore, the psychological problems of art can be approached with profit only by way of the psychology of the artist. Adopting this path of approach and taking into account the results of recent psycho-analytical research, I proceeded to consider the significance of the work of art as a product of artistic activity. In the present paper I shall pass from the psychology of the artist to the psychology of the beholder and consider the significance of the work of art as a source of aesthetic experience. The argument will, however, presuppose the conclusions of the earlier paper.

---

[1] Based on a paper read before the Scottish Branch of the British Psychological Society on 29 January 1938.

[2] The term 'beholder' is employed throughout this paper to signify "*the subject of aesthetic experience*". Although the argument is developed with special reference to pictorial forms of art, it is intended to apply to every form of art quite irrespective of the particular sense department involved. The term 'beholder' is now chosen in preference to the wider term 'percipient', which appeared in my former paper, to signify a percipient who is undergoing aesthetic experience.

II. Unwarranted assumptions in aesthetics

In my former paper I provided a background for the present attempt to determine the nature of aesthetic experience. The importance of such a background is not difficult to appreciate, when we consider the confusion which reigns in aesthetics, both as to what constitutes a work of art and as to what is implied in the conception of 'the beautiful'. Owing to its misleading associations, the term 'aesthetics' must itself be regarded with suspicion. The term is now employed almost exclusively in the secondary meaning for which Baumgarten is responsible (2), and which is defined in *The Shorter Oxford English Dictionary* as "The philosophy of taste, or of the perception of the beautiful". Such a definition provides an interesting example of the presence of unwarranted assumptions in aesthetics. According to the definition, it will be noticed, 'taste' is equated with 'the perception of the beautiful'. Herein lie two assumptions: (1) that 'the beautiful' is something objective, which has only to be perceived to give rise to aesthetic experience; and (2) that it is the necessary function of a work of art to be beautiful. These, however, are assumptions which can hardly be allowed to pass unchallenged; for they present us with the alternative of either denying the status of 'work of art' to all such professed products of artistic activity as are not universally acknowledged to be beautiful, or of so modifying the conception of 'the beautiful' as to enable products of artistic activity, about which there is no such consensus of opinion, to be admitted to this category. The fact is, of course, that there is no general agreement as to what is beautiful; and consequently standards of beauty provide no criterion of what constitutes a work of art. It is only when we seek our clue to the psychological significance of art in artistic activity (i.e. in the motives of the artist) that it becomes possible to establish a definite criterion by which to differentiate between what is a work of art and what is not. Such a criterion is to be found, however, in the conclusion reached in my previous paper that anything that is 'made for fun' is a work of art. This conclusion, together with the conclusion that the work of art represents a tribute of restitution paid by the artist's ego to his super-ego, provides a means of surmounting the difficulties inherent in prevalent conceptions in aesthetics.

III. THE BEAUTIFUL IN ART AND IN NATURE

Before we proceed further, it seems desirable to make some attempt to clarify the relationship between the beautiful in art and the beautiful in nature. The phenomena of nature, to which we commonly ascribe beauty, are not, of course, works of art in the ordinary sense. Nevertheless, it is interesting to note how generally the presence of beauty in nature has been cited in support of the teleological argument for the existence of God. The beautiful in nature always gives the impression of something that does not depend simply upon the operation of natural laws, something that is not absolutely necessary, something that is *added to* natural phenomena. In other words, it gives the impression of something that is there 'for fun'. Here we are reminded of the definition reached in my previous paper, that anything that is made for fun is a work of art. It would appear, accordingly, that, if the beholder does not regard every work of art as 'beautiful', he nevertheless tends to regard everything 'beautiful' as a work of art.

IV. THE WORK OF ART AS A 'FOUND OBJECT'

This conclusion provides a convenient starting-point for the investigation of one of the later developments of Surrealism—a development following the invention by Dali of what is described as 'the surrealist object'. (A special Exhibition of Surrealist objects was held in London in the autumn of 1937.) Surrealist objects are defined by David Gascoygne ((3), p. 108) as "objects functioning symbolically". These objects are for the most part constructed by the artist out of a strange medley of materials, which appear to him significant. One such object constructed by Dali consisted in a hairdresser's bust of a woman, from whose neck two Indian corn cobs were suspended by means of a collar, and upon whose head was laid a loaf of French bread surmounted by a metal inkstand representing Millet's *Angelus*. This Surrealist object is illustrated, under the title of "Retrospective Bust", by André Breton (4). According to Gascoygne, "These objects typify the more recent ideas of surrealism, which conceive super-reality as existing in the material world, objectively, as well as subjectively in the automatic thought of the unconscious" ((3), p. 109). In the light of this statement, it is easy to understand the next development of Surrealism, which consisted in an enlargement of the conception of 'the surrealist object' such as to enable it to include so-called 'found objects'. The 'found objects' of the Surrealist consist in external, and for the most part natural, objects, in the appearance of which he discovers a hidden symbolic significance, and which he therefore preserves and, so to speak, 'frames'. They represent to those who 'discover' them a union between the world of outer reality and the inner world of the dream; and in virtue of this fact they are regarded as possessing a certain 'super-reality', which obviously resides in their capacity to act as symbols of wish-fulfilment. Their 'discovery' is accompanied by an intense emotional experience, which is evidently of an aesthetic nature, since it leads their discoverer to regard them as essentially 'beautiful'.

The interesting thing about the Surrealist's emotional attitude to his 'found object' is that it represents *an intermediate point between the attitude of the artist and that of the beholder*. In so far as the artist is not responsible for the existence of the object, but simply 'finds' it, he may be said to play the role of a beholder as much as any member of the general public who 'finds' such an object in a Surrealist exhibition. On the other hand, in so far as he is the person who discovers a hitherto unrealized significance in the object and 'frames' it by isolating it from its environment, he may be regarded, in a quite intelligible sense, as the creator of the object. As regards the 'found object' itself, the interesting thing is that it is just a work of art and no more. It is really *a work of art approximating quantitatively to à zero value*—a minimal work of art.

V. THE HISTORICAL ORIGIN OF ART

Considered from the point of view just indicated, the 'found object' of the Surrealist may well throw some light upon the historical origin of art. As we are reminded by Herbert Read ((5), p. 13), it has been found that in many of the earliest cave-drawings the objects represented are drawn round some projection or other natural feature forming part of the cave-wall itself. Such natural features are comparable to the 'found objects' of the Surrealist. They evidently suggested to the prehistoric artist the figures (usually animals) which he felt impelled to portray. Like Surrealist objects, they were 'objects functioning symbolically' and represented a union between the outer world of reality and the inner world of wish fulfilment; but, instead of just framing them as the Surrealist frames his 'found objects', the prehistoric artist went a step further and made them the nucleus of a picture in the accepted sense. Considered from the quantitative point of view, the resulting work of art thus represented a considerably greater departure from a zero value than does the 'found object' of the Surrealist; and it was only a further departure from this zero value, when the prehistoric artist created a picture which required no natural object as its material nucleus.

VI. THE GENESIS OF THE WORK OF ART

(a) *The discovery of the object*

The relationship between the 'found object' and the work of art in the accepted sense may be illustrated with reference to a picture entitled "Visage Paranoiaque" by the Surrealist artist Dali ((6), Plate 30). Hung in the horizontal position, this picture is seen to represent a group of natives sitting before an African building covered with foliage in a desert setting; but, if the picture is rotated clockwise through 90°, the building with the natives sitting before it assumes the appearance of a human head. In the absence of knowledge regarding the circumstances which led to its composition, this picture appears little more than a practical joke played upon the public by the artist. It appears in a different light, however, when the following circumstances are taken into account. At a time when Dali was specially preoccupied with a certain period of Picasso's painting, he was sitting at his desk one day preparing to formulate his reflexions in writing; and, while looking through his papers for some notes, he came across a picture post-card, which had been sent to him some time previously. As his eye fell upon the post-card, he immediately interpreted it as representing a picture of a head by Picasso; and it was only later that he recognized it to be a photograph of an African village with natives sitting outside their huts (see Gascoygne (3), pp. 101 ff.). In the light of his original interpretation, however, the post-card had all the significance of a 'found object'; and this 'found object' provided the nucleus of the picture.

(b) *Wish fulfilment*

As a matter of fact, what really inspired Dali to paint the picture, to which reference has just been made, was the discovery of the process of discovery itself; and it is thus integral to the picture that the relationship of the 'found object' to the resulting work of art should be unusually obvious.[1]

This fact, together with the history of the composition, provides us with an exceptional opportunity to study the genesis of a work of art. The artist is confronted in the course of his daily experience with an object in which he discovers a symbolic significance. This significance is not inherent, however, in the real nature of the object. It is due entirely to the fact that the object happens to possess features which enable it to represent for the artist a fulfilment of his emotional needs—a 'wish fulfilment'. For the artist concerned, the discovery of such a significance in an object is tantamount to the discovery of a new object. The object thus acquires the quality of a 'found object'; and the artist seeks to preserve and perpetuate what he has found in the form of a work of art. At the same time, the 'found object' is not simply something that was there all along and is now revealed; for, since its significance is dependent upon its capacity to represent a wish fulfilment, it does not exist apart from the act of discovery. It is created by the discovery itself; and the discovery represents a creative act on the part of the artist. It is but a further elaboration of this creative act when the artist embodies his discovery in a formal work of art.

The process described may be taken to represent what happens in all artistic creation, Surrealist or otherwise. If the universality of the process is not obvious at first sight, this is perhaps chiefly due to the fact that it is not always easy to isolate the 'found object'. There need be no real difficulty, however, unless we interpret the 'found object' in an unduly restricted sense. The 'found object' may be anything in the external world, which serves the purpose of wish fulfilment for the artist. In the case of representational art, the 'found object' is the object depicted—not, of course, the object as it exists in outer reality, but the object in its emotional significance for the artist. In the case of expressionist art (and Surrealism, it may be noted, is really a development of expressionist art), the 'found object' may be derived from the combination of a number of objects, which are quite unrelated in outer reality, but which have an emotional relationship in the artist's mind. The same applies even in the case of abstract art; for, even when art is at its most abstract and is concerned solely with the arrangement of lines, planes, surfaces and colours, the composition represents a combination of emotionally related elements derived from the artist's experience of the external world.

---

[1] It is a common practice of Dali to portray in his pictures the processes underlying artistic creation, and in particular the process of symbolization.

(c) *The need for disguise*

In the light of my previous paper, we must, of course, regard the emotional needs, which determine the significance of the 'found object' for the artist, as predominantly unconscious. It is true that the Surrealist appears to be conscious to an unusual degree of the nature of the appeal made to him by his 'found objects'; but the case of the Surrealist is exceptional in this respect, that it is his conscious aim to effect a union between the world of reality and the world of the unconscious. As I pointed out in my previous paper, the comparative poverty of the art work in Surrealist works of art is evidence of a relative failure of repression. Consequently, it is only to be expected that in the case of Surrealist art the 'found object' should represent the demands of unconscious urges with an unusual poverty of disguise. In the case of most schools of art, however, the incidence of the 'found object' is determined by a greater need for disguise; and, consequently, it is usual for the ultimate significance of the 'found object' to be unconscious, although, of course, the fact of the object's appeal is consciously recognized. Even for the Surrealist, indeed, a certain amount of disguise appears to be necessary to enable any object to function as a 'found object'. At least this seems to be the meaning of André Breton, the leading exponent of Surrealism, when he says, "One's pleasure is always partly accounted for by the lack of resemblance between the desired object and the *discovery*" ((4), p. 42).

VII. THE NATURE OF AESTHETIC EXPERIENCE

Taking as our starting-point the attitude of the Surrealist to his 'found objects', we have now considered the discoverer of the 'found object' in the role of *artist*; and the conclusion which we have reached is that all artistic creation involves the discovery and perpetuation of a symbolically significant object, i.e. an object which symbolizes for the artist the fulfilment of unconscious emotional needs. If now we consider the discoverer of the 'found object', no longer as artist, but simply as *beholder*, we seem justified in concluding that aesthetic experience represents a specific emotional reaction, which occurs when a symbolically significant object is discovered in the external world.

Aesthetic experience may accordingly be defined as *the experience which occurs in the beholder when he discovers an object which functions for him symbolically as a means of satisfying his unconscious emotional needs.* Although, of course, the fundamental emotional needs of mankind are shared by all, nevertheless in the case of any given individual they are influenced by a considerable number of factors; and, in consequence, individuals vary both in the precise nature of their emotional needs and in the amount of disguise necessary to enable an object to function symbolically for them, i.e. to function as a 'found object' and not as an actual object. This fact enables us to appreciate the subjective character of the aesthetic judgement.

VIII. THE WORLD OF ART AS A WORLD OF 'FOUND OBJECTS'

The aesthetic character of the experience, which occurs in the beholder in the presence of a 'found object', remains unaffected, whether the 'found object' belongs to the world of nature or belongs to the world of art. The world of art differs, however, from the world of nature in that it is composed specifically of 'found objects', whereas in the world of nature the 'found object' is only incidental. It does not follow, of course, that every work of art will appear beautiful to everyone. Indeed, we have already seen that the contrary is the case. Nevertheless, we must assume that every work of art must appear beautiful at some time to someone—if only to the artist who creates it.

If the world of art represents a world of 'found objects', the artist is evidently a person who is specially good at discovering such objects. A capacity of this nature is by no means the sole qualification of the artist, however; for, in the mere discovery of a symbolically significant object, the artist is as yet little more than a beholder, although he may have the capacity for discovery—the gift of 'vision'—in an unusual degree. The beholder only becomes an artist when he isolates or frames the significant object and gives it a certain permanence by, so to speak, converting it into a monument. The task of so doing constitutes artistic creation.

The most primitive form of artistic creation is represented by the 'found objects' of the Surrealist. In this case the 'found object' is merely framed more or less as it is found. A slightly more complex product of artistic creation is represented by such 'surrealist objects' as are constructed out of a variety of actual objects. It is but a further step along the path of artistic creation when symbolically significant objects come to be represented in pictures and sculptures. Once this stage is reached, the range of 'found objects' which can be so represented becomes, of course, enormously enlarged. The whole gamut of natural phenomena is placed at the disposal of the artist; and the possibilities of artistic creation are thus made more commensurate with the range of the artist's vision.[1]

The role of the beholder of a work of art is to discover for himself the objects which the artist has already discovered and isolated for him; and it is for this reason that the element of *surprise* is integral to the aesthetic experience. It is to this element that Herbert Read refers in *The Meaning of Art*, when, speaking of the frame of mind in which the beholder ought to approach a picture, he says "All that is necessary is that he should have a perfectly *open* mind....He just walks round a corner, thinking of nothing in particular, and comes to a standstill before this object.... Many theories have been invented to explain the workings of the mind in such a situation, but most of them err, in my opinion, by overlooking the instantaneity of the event" ((7), pp. 17 and 18). The importance of the 'open mind', to which Herbert Read refers, resides, of course, in the fact that the open mind is a precondition of discovery; and the 'instantaneity of the event' is characteristic of the act of discovery itself. In walking round a corner and finding himself confronted with a work of art, the beholder is thus placed in a position to 'discover' a symbolically significant object. It is true that the object is already a 'found object' in that it represents what its creator (the artist) has already discovered. Nevertheless, it is a new discovery for the beholder. In making his discovery, the beholder shares the experience of the artist who made the discovery in the first instance; and, by identifying himself with that artist, he shares the satisfactions of artistic creation himself.

IX. THE FAILURE OF AESTHETIC EXPERIENCE

(a) *Over-symbolization and under-symbolization*

It is always possible, of course, for the beholder of a work of art to fail to identify himself with the artist who created it. When this happens, as it frequently does, the work of art fails to present itself to the beholder as a 'found object' and no discovery takes place. Such a phenomenon may be simply due to a lack of technique on the part of the artist—as when, for example, a painter lacks sufficient mastery over his material to know how to produce the effects which he desires. It may occur, nevertheless, even when the technique is good. In this connexion I am reminded of a visit which I paid to the Royal Academy of 1936. As I went round the gallery, it so happened that I kept step with a clergyman; and I noticed that, each time he came to a picture by Stanley Spencer, he turned to his companion and expressed an unfavourable verdict in unmeasured terms. His final outburst was to this effect: "Did you ever see anything so dreadful? This isn't really art at all. And the worst of it is that the man can paint!" Obviously Stanley Spencer's pictures did not present themselves to this particular beholder as 'found objects'—and this in spite of the fact that he could not help recognizing the artist's mastery of technique.

---

[1] It seems permissible to speculate that the psychological processes, which led primitive man to immortalize 'found objects' in the form of works of art, may also have led him to immortalize such objects in the form of specific sounds, and that it is to this fact that we must look for an explanation of the origin of *language*. If this speculation is correct, then language must be regarded, not only as providing a medium for literary art, but as having originally constituted a form of art in itself. Indeed, in much of what is known as "Modern Verse" language would appear to have reverted to just such a status.

When such a situation arises, the reason why the work of art does not appear in the light of a 'found object' is that there is a disparity between the total emotional needs of the artist and those of the beholder—'*total* emotional needs' because it is a question, not only of the expression of repressed urges, but also of the demands of the super-ego and of that need for 'restitution', to which reference was made in my previous paper. This being so, the failure of a work of art to appeal to the beholder as a 'found object' may be due to either of two causes:

(1) It may be that the censorship of the artist's super-ego is so rigorous, and necessitates such an elaborate disguise of the urges expressed, that the work of art is deprived of almost all symbolic significance for the beholder. In such a case the work of art makes no emotional appeal. It falls flat. It conveys the impression of sterility. It provokes the criticism that the artist 'has nothing to say'. The beholder looks at it and makes no 'discovery'. Such is the impression created upon the artistically educated of the present generation by such Victorian works of art as the pictures of Lord Leighton.

(2) The inability of a work of art to present itself as a 'found object' may occur for a totally opposite reason. So far from being due to excessive censorship on the part of the artist's super-ego, it may be due to a relative failure of such censorship. When the artist's super-ego is weak and his repressed urges are really 'urgent', these urgent urges express themselves in the work of art with a minimum of disguise. The effect of such a work of art is to provoke an excessive emotional reaction in any beholder whose super-ego is more exacting than that of the artist, and whose need for the disguise of repressed urges is greater. So far from saying too little, such a work of art says too much. It says more than the beholder's super-ego will tolerate. Consequently, under the pressure of

impending anxiety, the beholder rejects it with feelings of disgust or indignation. In the presence of such a rejection, of course, the work of art cannot present itself as a 'found object'. The experience of 'discovery' is precluded because the object represented actually fails to convey the impression of symbolic significance to the beholder; and the reason of this failure is that the object is not sufficiently symbolic to pass the censorship of the super-ego. It is, so to speak, too like the real thing. This applies, so far as most people are concerned, to Surrealist works of art in general.

(b) *Form without content, and content without form*

It is to be noted that, in both cases instanced, the reason why the work of art fails to present itself as a 'found object' is that it lacks symbolic significance for the beholder concerned. The reason for this differs, however, in the two cases. The work of art lacks symbolic significance in the former case owing to an *over-elaboration* of disguise, which precludes any appeal to the repressed urges, in the latter case owing to an *inadequacy* of disguise, which leaves the requirements of the super-ego unsatisfied. It lacks symbolic significance in the one case because it is so over-symbolized that its symbolic reference becomes completely obscured, and in the other case because it is so under-symbolized that it has not reached the stage of being symbolic at all. To express the same thing in other terms, we may say that in the former case the symbolism has form without content, and in the latter case it has content without form. It is only in such a reference, at any rate, that any useful meaning can be attached to the contrasted terms of 'form' and 'content', which have long been a source of confusion in aesthetics.

(c) *The optimum of symbolic significance*

At this point an observation of considerable significance may be recorded. Whilst there is a marked difference in the emotional reactions produced in the observer by the over-symbolized and the under-symbolized works of art respectively, yet both tend to provoke a common criticism. Both are equally liable to call forth the comment, "This isn't really art at all". This fact is important because it indicates that, for any given beholder, the work of art depends for aesthetic value upon the degree to which it approximates to an *optimum of symbolic significance* lying at an intermediate point on a scale between over-symbolization and under-symbolization. In other words, a work of art only acquires an optimum of aesthetic value for the beholder in so far as it enables both his repressed urges and the demands of his super-ego to obtain a maximum of satisfaction.

X. THE PRINCIPLE OF RESTITUTION

(a) *The work of art as a 'restored object'*

The duality of satisfaction which we have just found to underlie aesthetic experience is now seen to correspond to the duality of satisfaction provided by artistic activity; for, as was pointed out in my previous paper, artistic activity performs the double function of providing a means of expression for the repressed urges of the artist and of simultaneously enabling his ego to pay a tribute to the supremacy of his super-ego. It was also pointed out in my previous paper that, in so far as the work of art consists in a tribute paid by the artist's ego to his super-ego, it essentially represents a means of restitution, whereby his ego makes atonement to his super-ego for the destruction implied in the presence of repressed destructive impulses. In pursuance of the analogy between the emotional needs of the beholder and those of the artist, we may now infer that the aesthetic appeal of a work of art will depend upon its capacity to represent a restitution for the beholder also. The aesthetic experience is thus seen to be an experience that occurs when a person finds himself confronted with an object which presents itself to him, not simply as a 'found object', but also as a *restored object*.

Now that a work of art is seen to depend for aesthetic value upon its capacity to function as a 'restored object' for the beholder, we can understand better how it comes about that aesthetic experience should be compromised in the presence of both under-symbolization and over-symbolization. The under-symbolized work of art fails to produce the effect of restitution, because the impression which it gives is one of more or less unmitigated destruction and, consequently, the demands of the super-ego remain unsatisfied. On the other hand, restitution is meaningless apart from the presupposition of destruction; and, consequently, over-symbolization precludes the effect of restitution by excluding the impression of destruction too rigorously.

(b) *The optimal synthesis*

Here we are at once reminded of the central conception of Hegelian philosophy—a conception summarized by Edward Caird in the statement that "the highest unity is to be reached only through the full development and reconciliation of the deepest and widest antagonism" ((8), p. 42). Taking advantage of this philosophic conception for descriptive purposes, we may say that aesthetic experience depends upon the resolution of an antinomy created by the simultaneous operation of the libido (the life principle) and the destructive urges (the death principle). In the resolution of this antinomy the demands of the libido may be said to constitute the thesis, the pressure of the destructive urges the antithesis, and restitution the synthesis. What conditions aesthetic experience within the field of art is the capacity of the work of art to represent such a synthesis for the beholder. For a complete synthesis to occur, there must be an optimum proportion in the relationship between thesis and antithesis. It is only in so far as such an optimum relationship exists that the work of art becomes a 'restored object'. In a work of art characterized either by under-symbolization or by over-symbolization the proper relationship is absent, and the aesthetic experience is accordingly compromised.

XI. AESTHETIC THEORY

It is to the dependence of aesthetic experience upon the satisfaction of a need for restitution in the beholder that we must look for the rationale of that type of aesthetic theory which seeks the clue to 'beauty' in a certain 'perfection' or completeness in the object. This type of theory has, of course, played an enormous part in the history of aesthetics. It is true that it has had fewer advocates in recent years, as psychologists, anthropologists, art critics and even artists themselves have increasingly invaded the field of aesthetics, formerly so largely monopolized by philosophers. Nevertheless it has always made a special appeal to the more philosophically minded; and it has undoubtedly dominated the field of aesthetic theory in the past. Its most concise formulation is perhaps that of Aristotle: "The essential characters composing beauty are order, symmetry and definiteness" (9). Such a formulation is representative of a widespread tendency in aesthetics to regard the 'beauty' of a work of art as residing in its capacity to convey the impression of something *intact*, *whole*, *complete* or *perfect*; and the presence of this tendency may be regarded as confirming the correctness of the view that aesthetic experience depends upon the satisfaction of a need for restitution in the beholder.

In the history of aesthetics there is evidence, however, of a contrasting line of thought—a line of thought perhaps best epitomized in Croce's claim that "expression and beauty are not two ideas but one" (10). Expressionist theories of aesthetics seem to make a special appeal to the modern mind, characterized as it is by a tendency to question time-honoured values. That they are no novelty, however, may be illustrated by reference to Aristotle himself; for, in spite of his 'perfectionist' theory of aesthetics in general, we find Aristotle ascribing the appeal of tragedy to its capacity to effect a catharsis of the emotions of pity and fear—an essentially expressionist interpretation (11). Considered in isolation, such a theory is hopelessly inadequate to explain the aesthetic value of tragedy, since the emotions of pity and fear in themselves tend to compromise aesthetic values rather than to enhance them—the reason being, of course, that pity and fear imply a sense of uncontrollable destruction at work. The tragedy which simply releases pity and fear is thus comparable to the under-symbolized picture. Its effect is to produce a state of distress or horror in the beholder, and to provoke a reaction of rejection on his part. On the other hand, a tragedy which fails to release pity and fear in the beholder is rejected as emotionally unconvincing. It appears trite or ludicrous, because it fails to convey any impression of destruction at all, thus excluding the possibility of restitution. Such a tragedy is comparable to the over-symbolized picture.

It becomes evident, therefore, that it is only when a tragedy complies with perfectionist as well as expressionist standards that it acquires aesthetic value. To impress the beholder as in any sense 'beautiful', a tragedy must not only release the emotions of pity and fear, but at the same time produce an impression of completeness and perfection—an impression of '*the integrity of the object*'. The fact is that the expressionist and perfectionist theories of aesthetics each represent a half-truth—the whole truth being that a work of art can only appear 'beautiful' to the beholder in proportion as it satisfies the particular conditions demanded by his need for restitution. To satisfy these conditions (i.e. to appear beautiful), the work of art must be able above all to produce in the beholder an impression of the 'integrity of the object'; but, in order to do so, it must at the same time provide a release for the emotions which imply the destruction of the object. Otherwise the conditions of restitution remain unsatisfied; and aesthetic experience is accordingly precluded.

REFERENCES

(1) FAIRBAIRN, W. R. D. (1938). "Prolegomena to a psychology of art." This *Journal*, XXVIII, 288–303.

(2) BAUMGARTEN, F. (1750). *Aesthetics*, § 14. Quoted by CARRITT, E. F. (1931). *Philosophies of Beauty*, p. 48. Oxford Univ. Press.

(3) GASCOYGNE, DAVID (1936). *A Short Survey of Surrealism.* London: Cobden Sanderson.

(4) BRETON, ANDRÉ (1936). *What is Surrealism?* London: Faber and Faber.

(5) READ, HERBERT (1937). *Art and Society.* London: Heinemann.

(6) DALI, SALVADOR (1935). *La Conquête de l'Irrationnel.* Paris: Éditions Surréalistes.

(7) READ, HERBERT (1931). *The Meaning of Art.* London: Faber and Faber.

(8) CAIRD, EDWARD (1901). *Hegel.* Edinburgh: Blackwood.

(9) ARISTOTLE. *Metaphysics*, XII, 3. Quoted by CARRITT, E. F. (1931). *Philosophies of Beauty*, p. 36. Oxford Univ. Press.

(10) CROCE, B. *A Breviary of Aesthetics.* Quoted by CARRITT, E. F. (1931). *Philosophies of Beauty*, p. 243. Oxford Univ. Press.

11) ARISTOTLE. *Poetics*, VI.

PROPOSED EXPLANATION OF INDIVIDUAL DIFFERENCES OF ABILITY BY "SAMPLING"

By C. SPEARMAN

I. *A long controversy* (p. 182).
II. *Alleged hierarchical "tendencies"* (pp. 182–183).
III. *Alleged bias of selection* (pp. 183–184).
IV. *"Main" Two-Factor theory* (pp. 184–185).
V. *Sub-theories: Energy versus Sampling* (pp. 185–186).
VI. *Versions of "Sampling" repudiated by the Samplists* (pp. 186–187).
VII. *Version now advocated* (pp. 187–188).
VIII. *Objections* (pp. 188–189).
IX. *Psychological service* (pp. 189–190).
X. *Conclusion* (p. 190).
References (p. 191).

I. A LONG CONTROVERSY

For several years, the theory of ability known as that of "Two Factors" has been energetically opposed by that of "Sampling". In 1931 the present writer found himself obliged to consider this doctrine much more systematically than before. But the result was still adverse (1). Later on, however, notably by Thomson in 1935, the sampling theory was emphatically reaffirmed (2).

The following article aims at making the issues clearer. One special motive for so doing is that the time seems to have now arrived, if not indeed for eliminating all the doctrinal discord, at least for rendering its scope much narrower.

Another reason for my present intervention is that Thomson, although very kind and ample in his general allusions to myself, appears to have overlooked just my aforesaid writing, which explicitly and systematically dealt with the subject. Not only does this work find no place in his references, but also the points raised by it get from him little or no reply.

II. ALLEGED HIERARCHICAL "TENDENCIES"

For the purposes of clarification, let me first of all try to define better certain current expressions which have been peculiarly misleading. Our whole Two-Factor doctrine has been based on the theorem that—under certain statable conditions—the intercorrelations of any set of ability-scores tend to a regular arrangement. This is designated as hierarchical and is proved by the fact of certain quantities, called tetrad differences, being equal to zero. But against this position, Thomson and his followers urge that a mere "tendency" to zero tetrad differences arises necessarily from, and therefore is more properly explained by, the bare action of Random Sampling.

His and our usages of the word tendency, however, lie very far asunder. In the language of the proponents of Sampling, the tetrad differences are said to "tend" to zero when such is their average value (or their most probable one).[1] But *this* zero automatically ensues from the bare fact of the tetrad differences in question being symmetrically distributed with *opposite signs*. It is really devoid of significance for the Two-Factor theory, or indeed for that of Sampling.

In the language of the Factorists, quite otherwise, the tetrad differences said to tend to be zero are those which are taken *without regard to signs*, or absolutely. And what the Factorists have here meant by "tendency" is that the values are not indeed perfectly zero as actually observed, but become so when all allowances have been made which are definitely prescribed by the theory (as, for instance, the proper compensations for probable errors, or those for duly verified group factors). Any zero or hierarchy still more perfect than this would not improve the evidence for the theory, but only mar it.

As evidence for such a theoretical tendency of the absolute values to perfect zero, we may quote especially the researches of Stephenson and those of Brown. The former author, using 1000 subjects, found the observed median deviation from zero to be no more than 0·008, which could be accounted for exactly by the error of sampling, since the probable value of this was practically the same, namely 0·009 (4).

More detailed still was the statistical agreement of observation with theory reached next year by Brown (5, 6).

III. ALLEGED BIAS OF SELECTION

We pass on to another and even graver trouble. The Samplists have charged the Factorists with biased procedure. We are depicted as "selecting" or "picking" those tests and correlations which are favourable to the hierarchy and "ignoring" or "trimming" the remainder (see (2)). In truth, however, far from favouring the cases that conform to the hierarchy, we have been just as much concerned to discover the cases that fail to conform. In fact, our studies have been mainly devoted to determine the position of the *border line.*

---

[1] See, for example, Thomson ((3), pp. 69–70); also repeatedly (2).

On certain minor points, no doubt, our theoretical anticipations have sometimes encountered disappointment. But very soon these seeming departures from the theory have turned out to be really extensions of it. Thus on one occasion they could be explained by, and so led to the discovery of, a further spatial factor, which has since been abundantly verified (7). In much the same way, long before, a seeming discrepancy of the experimental results of Hargreaves had led him to the detection of the subsequently corroborated factor of "fluency" (8). Such discoveries are analogous to that which brought astronomers, from observing disturbances in the motion of Uranus, to disclosing the existence of Neptune. Another physical instance was when an observed excess of weight in a chemical analysis guided to the detection of a new element (lithia). All such cases exemplify Mills' familiar scientific method of "residues".

IV. MAIN "TWO-FACTOR" THEORY

Let us now turn to where the Samplists and the Factorists do admittedly come into line. This surely is a point deserving more general and consistent recognition. Thomson himself writes generously enough:

> I have never denied the possibility of Professor Spearman's Theory of Factors, as some seem to think....I think that his system is an admirable *description* of correlation data (see (2), p. 64).

But he goes on to add that this system does not give *causes*. The factors, he declares, are no "entities" or "organs"; they have no "real existence", any more than has a mere average, or a standard deviation. And if, as is the impression conveyed, he implies that such causes form an integral part of our theory of Two Factors, he is gravely misrepresenting us. A "factor" with us has only its ordinary meaning as given in the dictionary. If meaningful (as opposed to merely statistical) it is taken to be: "one of the circumstances, facts or influences which produce a result".

In this reference we Factorists may fairly complain that the opponents of our view persistently overlook what has always been one of its most fundamental features. This consists in the distinction made by us between what we have called respectively our "main" theory and the supplementary "sub-theories" (9).

The theory includes a number of theorems. The chief of these may be formulated as follows. Every ability score can be divided into two parts called respectively "general" and "specific". These two are uncorrelated. Different specific abilities, on the other hand, are uncorrelated with each other when, and only when, their mutual likeness (or other tie) sinks below a certain critical point. With any greater likeness, these abilities do become correlated. And the production of correlation in such a way is responsible for the so-called "group factors". The primary development of this theory consists in proving the existence of these critical points and determining their positions. But secondarily, many further single theorems are introduced for the purpose of building up a complete theory of greatest attainable service. Most of them are collected in the Appendix to my *Abilities of Man*. All alike deal only with factors as quantitative influences, *not* as "organs". And so far, the Factorist and the Samplist views would appear to be, after all, in general agreement.

Whilst on this subject, however, a few words may be said on some other points that seem to have been embarrassing. One is a distinction that has been stressed between the "theory" and the "method" of Two Factors. Really, both words are used by us in their natural senses: "theory" means a general way of looking at things: "method" is a procedure for attaining to any object. Both depend on the same set of separate Two-Factor theorems.

But why "two" factors and not rather one, or three, it has been repeatedly protested? To this we would reply that in our view we can divide up an ability into as many factors as we please, all of them equally true. But some divisions will have much greater general importance than others. And the division which is most important of all, it seems to us, is into two parts, the general factor and the specific factor. Next in interest comes the subdivision of the specific factor into correct value and mere random error. But, as said, to the possibility of such divisions there is no finite limit.

So much for the main theory of Two Factors. Elsewhere, I have given chapter and verse to show that there has been no essential change in it from its first enunciation in 1904. And it might be fairer if those opponents who state or imply that I have ever made any radical alterations would try to substantiate this charge from my own words; not, as Thomson and others appear to do, only by passing it on from one to another (10).

V. SUB-THEORIES: ENERGY VERSUS SAMPLING

Supplementary to all this, our rather confidently advocated main theory which envisages mere quantities, have come the more or less speculatively advanced sub-theories, by which such quantities may be interpreted and *explained*.

Even as regards these sub-theories, however, Thomson and I really go a long way hand in hand. In his most emphatic expression of his doctrine he has written as follows:

> My own view is that the mind is one extremely complex organism which, at the time when it comes to attempt any of the tests used in the production of these tables of correlations, draws its ability to succeed in any one of them from a multitude of sources, some inherited genes for this or for that, some points of training or of experience during education and during life.

And on these words I proceeded to make the following comment:

> In itself this view is irreproachable. Indeed this or something like it has long been agreed upon by everybody. The real problem is not as to whether some such situation of multiple elements exists or not, but as to how it can best be harmonized with another known situation (the correlational hierarchy)....One attempt—hypothetical indeed, but fitting admirably with an immense number of known facts and contradicted by none of these—is to regard the genes as dominantly responsible for producing a general factor of power, whereas the "points of training or of experience" are taken to be the main contributors to the specific factor ((1), p. 158).

Accordingly, he and I alike conceive the correlations to derive *ultimately* from "a multitude of sources". We only disagree as to the manner that we respectively think of the derivation as occurring.

For my part, I get most help from supposing that the multitudinous particular small causes act by way of all contributing to one and the same general "power", "energy", "potential", or, in particular, Lashley's "mass action".

Thomson, on the other side, thinks he can get from the original multitude of small causes to the observed hierarchy by merely taking "samples" of the former and "leaving the rest to chance". This, he said, "will give hierarchical arrangement in the overwhelming majority of cases": "hierarchical order will arise unless we take pains to suppress it" ((11), p. 341).

VI. VERSIONS OF "SAMPLING" REPUDIATED BY THE SAMPLISTS

However, this notion of Sampling has proved unexpectedly troublesome. Already in 1914 I had myself brought forward what seemed to me the two most natural forms of it and had shown that they could *not* produce, or even conform with, the observed hierarchical arrangement.[1] But Thomson summarily disposed of this earlier work by remarking that we had taken "an entirely different meaning of the word Sampling" ((2), p. 66).

---

[1] See especially (12); of cardinal importance here are various works of J. C. M. Garnett.

Undaunted, Brown and Stephenson tried to improve matters by accepting the version of Sampling used in the admirable research of Blackie, who was a student of Thomson and had worked in consultation with him. But the result was again that the Sampling theory broke down (13, 14). And once more Thomson disposes of the matter by saying that "the adoption of these equations (of Blackie) means giving up entirely what I had meant by Sampling" ((2), pp. 66, 70).

Altogether, Thomson seems to have moved very far from his original attempt to explain the hierarchy by *any* chance or random arrangement of elements. The hierarchy now appears to be only explicable by a sampling of some distinctive kind not easy to apprehend.

Small wonder, then, if Brown wrote resignedly that the comparison of the two theories at issue:

> ...is handicapped by the inadequacy of the formulae hitherto brought forward in connection with the Sampling Theory ((15), p. 398).

VII. VERSION NOW ADVOCATED

With the further aid of his present exposition, however, let us make yet another attempt to capture this elusive notion of "Sampling". Originally, Thomson wrote as follows:

> Let us suppose that the mind, in carrying out any activity such as a mental test, has two levels at which it can operate. The elements of activity at the lower level are entirely specific: but those at the higher level are such that they may come in play in more than one kind of activity, in more than one mental test... (11).

Next year he seems to have dropped the two levels, but otherwise to maintain as strongly as ever that

> Hierarchical order is the natural order to be found among any set of correlation coefficients, *however the correlations may be caused*, provided only that the bonds which cause the correlations are left, as we say, to chance: that is, provided the phenomenon is of that complex nature which, in default of analysis, we call *random* (16).

More definitely he wrote seven years later:

> Let there be a universe of A elements, all-or-none in nature, and quantitatively equal in their effects, of which universe a_1, a_2, a_3, and a_4, are samples of those numbers respectively of such elements. To clarify ideas (though the illustration will limit and coarsen the possibilities), A may be thought of as the totality of neurones in the cortex, a_1 the *neurones concerned in the first activity*, and so on (3).

Finally, eight years later still, he takes each variate (test-score) to be:

> The straight unweighted sum of its components or elements, which were a sample taken from the pool of all such elements (2).

This doctrine he now expresses in the following equation, one however which is not informative, since it is common to all theories on the subject. It is, as it were, but the canvas on which has still to be painted the picture itself:

$$z_i = m_{i1} e_{x1} + m_{i2} e_{x2} + \dots + m_{iN} e_{xN},$$

where z_i is the score of the individual z at the test i, the e's are the factors entering into this score and the m's are their respective coefficients, weights, or loadings. Turning to the qualifications of this equation, which do distinguish and characterize the theory of Thomson, these seem to include especially the following assumptions. If they are not really his, we should be very glad to have from him a revision of them.

Assumptions of Sampling Theory

(1) The Universe of N elements is extremely large. This hypothesis is implied in the suggestion that the e's may represent a person's neurones.

(2) The N m's are restricted to the values one or zero.

(3) The n m's entering into any test constitute a "sample" of all the m's.

(4) The n e's of any individual in the test constitute a "sample" of all his e's.

(5) Each test has its own arbitrarily defined "richness"; that is to say, the proportion of m's which are not zero.

VIII. OBJECTIONS

Now, to these fundamental assumptions of the sampling theory, there would appear to be the following objections:

(*a*) A sample, whatever else it may mean, consists certainly of a number of units drawn from the chosen universe independently of one another. And statisticians know that such a sample is extremely difficult to find in any field. Least of all seems such an independence to be reasonably expected between the elements entering into such a highly organized thing as a mental test. One must think rather that such a sample of independent elements as is indicated in assumptions (3) and (4) could only produce nonsense.

(*b*) If the said assumptions did hold, then they would not merely cause—as Thomson says—a "tendency" to hierarchical arrangement. Instead, since N and therefore n (if there is to be correlation) are extremely large, the hierarchy would needs be as good as perfect. This again is in conflict with experience.

(*c*) The third of the preceding assumptions appears to preclude the fifth. If the *m*'s entering into any test were really a very large sample, then the "richness" could not possibly vary from test to test; for by the law of sampling $s_N^2 = pq/N$, approximately $=0$, where p and q are the respective chances of an m being one or zero, whilst s is the standard deviation of the proportion of the *m*'s that are = one.

(*d*) The preceding conditions would seem to produce many further results incompatible with actual observations. For instance, by virtue of (1) and (4) an individual would have the same standardized score at every test. Again, all the correlations would be equal.

On the whole, then, Thomson's "Sampling" theory of abilities, in its present state of development, seems to be definitely inacceptable. As may be seen, the chief troubles arise from the hypothesis that N is large. But if, instead, it be made small, then other no less fatal difficulties supervene. The whole plausibility of the sampling theory has only originated from equivocating between the one hypothesis and the other (1).

Nevertheless, the importance of this result should not be overestimated. After all, we have only been dealing with very speculative hypotheses, about whose truth or untruth there appears to be in any case no finality. Each is only cherished for as long as it continues to be scientifically helpful.

IX. PSYCHOLOGICAL SERVICE

But here we come upon a more serious danger. How far do, in fact, the two rival hypotheses fulfil their essential mission of stimulating scientific progress?

Consider first, the theory of Two Factors, beginning with the general factor "*G*". Surely, we have been fruitfully led to make numerous inquiries about its nature and function. Does the influence of this factor increase in proportion as the ability at issue is "noegenetic"? At what age does it maturate? To what extent is it acquired by experience? Or take any specific factor "*S*"; say, the "verbal" one. We have been brought to ask what mental processes this factor involves.

We are urged to enquire how far it enters into children's difficulties in learning to read. Again, in what fashion, if at all, is it connected with pathological disturbances of speech? Answers to these and innumerable analogous questions seem to be cogently needed for all conduct of life; for education, medicine, industry, sociology, genetics, and so on without limit.

Now in respect of these momentous problems, the Factorists claim to have achieved steadily progressive information year by year, as befits genuine science.

But where, on the other hand, does the work of Thomson and his followers come into the picture? For my own part, with all good will, I cannot answer this question satisfactorily. All the other though conflicting schools of research—as those of Kelley, Thurstone, Holzinger, Brown and Stephenson—would appear at least to be helpfully checking and counter-checking one another, so as gradually to build up one great agreed edifice. Whereas Thomson's elaborate excursions instigated by the theory of sampling—in themselves generally most ingenious, and even irreproachably correct—would appear to have little or no contact with any of the vital matters of psychological significance.

X. Conclusion

Altogether, our examination of the "Sampling" theory of ability has led to an unfavourable issue. According to our considerations, it would seem to have been obscure in expression; to contradict admitted facts; to tempt its proponents into misrepresenting the rival Two-Factor theory; and, worst of all, to divert their own work from contact with actual mental experience, and therefore from benefit to psychological science.

Still, this Sampling theory is after all only hypothetical. In consequence, it can readily be discarded without in any way disturbing the actual facts, either mathematical or observational. And on such release from it, the doctrine urged by the Samplists turns out to be no other than that which has always constituted the "main" theory of the Factorists themselves. In this fashion, the way appears wide open to reconciliation and collaboration.

REFERENCES

(1) Spearman, C. (1931). *Brit. J. educ. Psychol.* i, 140.
(2) Thomson, G. H. (1935). This *Journal*, xxvi, 63.
(3) —— (1927). This *Journal*, xviii, 68.
(4) Stephenson, W. (1931). *J. educ. Psychol.* xxiii, 185.
(5) Brown, W. (1932). This *Journal*, xxiii, 171.
(6) —— (1932). *Nature, Lond.*, cxxx, 588.
(7) Stephenson, W. (1935). This *Journal*, xxv, 197.
(8) Hargreaves, H. L. (1927). *Brit. J. Psychol. Monog. Suppl.* iii.
(9) Spearman, C. (1920). *Psychol. Rev.* xxvii, 159.
(10) —— (1933). *J. educ. Psychol.* xxiv, 591.
(11) Thomson, G. H. (1919). This *Journal*, ix, 337.
(12) Spearman, C. (1928). This *Journal*, xviii, 249.
(13) Brown, W. and Stephenson, W. (1932). This *Journal*, xxiii, 352.
(14) Brown, W. (1934). *Nature, Lond.*, cxxxiii, 724.
(15) —— (1935). This *Journal*, xxv, 395.
(16) Thomson, G. H. (1920). This *Journal*, x, 319.

(*Manuscript received* 31 *May* 1938)

PUBLICATIONS RECENTLY RECEIVED

Driver Testing Results. By HARRY R. DE SILVA and THEODORE W. FORBES. Harvard Traffic Bureau. 1937. Pp. vi+98.

This rather slipshod report presents the results of testing 1776 motor-car drivers drawn from five American cities. The battery of tests used comprised a Standard Vigilance Test (designed to test speed of reaction, manipulative ability and 'Vigilance'), a Speed Estimation Test, a Glare Test, a Miniature Highway Driving Test (designed to test simultaneously complex judgments of speed and timing, and accuracy of manipulation of car controls), a test of Peripheral Vision, an Auditory Test, a test for Colour Vision and a test of Visual Acuity. In addition each subject was asked to answer a questionnaire on personal characteristics, including questions on length of driving experience, customary road speed, number of years of academic education, parentage, civil status, pigmentation, etc.

The drivers tested were divided into 'Volunteer or Average Drivers' (male and female) and 'Accident Repeaters'. No definition of these two categories is given and it is not at all clear on what basis they were selected. For instance, in a section on illustrative cases of drivers who have had accidents, it is said that "many of the drivers have only one or two accidents on their record, whereas they might easily have others and perhaps will have others in the future", whilst two pages later it says, apropos of local differences in test scores, "This interpretation assumes that the volunteer group is a fair average sample in each case." One wonders if the first set of drivers with only one or two accidents on their records in three years were called 'Repeaters' and whether steps were taken to obtain and examine accident records for the 'Volunteer' group. The comparisons between the test scores of these two groups would have gained much more weight if evidence had been given to show that the 'Average' groups of drivers really could be used as controls to the 'Repeater' groups.

Assuming, however, that the comparison between the two groups is legitimate, one reads that the accident repeaters show "a consistent and rather general deficit in the abilities represented in the test battery". On examination of the statistical tables at the end of the report it is apparent that the consistency is not very marked, as the differences between test means are given as statistically significant in two out of four cities, one city showing insignificant results in every case. Of this city the authors state: "Table 4 (City E) is at variance with the other three, probably due to lack of development of the tests and less training of the examiner at the time of testing." This may be so, but no such reason is advanced in the case of City C, where only four out of eleven comparisons can be regarded as significant on the basis of these tables. However, it appears on the whole that the accident repeaters are poorer in certain of the tests than are the 'average' drivers, a result which agrees with work done by other investigators in other countries.

Numerous other comparisons of greater or lesser importance are made. Stated briefly, the authors find that the abilities tested by certain of the tests decrease with age, the tendency beginning to be most evident at between thirty and thirty-five years of age: that in general, commercial drivers were somewhat above average at the tests: that in general, women drivers are poorer than men in tests involving muscular quickness and accuracy. Little of definite value seemed to emerge from an examination of outstanding personal characteristics.

Certain significant differences between the test performances of 'volunteer' drivers drawn from different cities were found, and the authors state that "differences obtained are interpreted to mean that certain factors such as congestion and the type of street layout and other driving conditions in certain areas call for different degrees of skill for successful driving in certain cities". This interpretation is exceedingly doubtful since it assumes an almost complete identity between driving ability and the test used, and although the authors show fairly convincingly that there is a relation between the tests and accident incidence, it cannot be assumed without evidence that there is a perfect association between driving ability and accidents sustained. The authors throughout seem to assume that a driver who has not sustained any accidents in the particular period of observation chosen is *ipso facto* a good driver and that a driver who has had an accident is a bad one. This is, of course, allowable as a definition, but without a qualitative examination of the individual accidents it is hardly warrantable to assume an identity between driving ability and number of accidents. In any case it seems unlikely that the psychological qualities tested should really differ from city to city on account of differences in traffic conditions.

A most important statement is made on page 49 of the report. Under the heading 'Re-education Value', the authors make the following statements. "The largest group tested consisted of 470 male accident repeaters. This group was made the subject of a follow-up study on accident records and proved to have a 90% reduction in accidents during the following six months, as compared to a similar group of accident repeaters who were given no tests or education. This, therefore, constitutes evidence that the tests have very definite and specific value in the re-education of accident repeaters and poor drivers." One searches throughout the report in vain for the evidence on which this important point is based. Obviously, if it is true, the possibility of re-educating accident-prone drivers is extremely important as a practical method of reducing accident incidence and one would have liked to have seen the evidence for the striking example of its efficiency which is given. All that is apparent, however, is that the group referred to is City E, and the authors elsewhere throw doubts on the reliability of the tests given to this group (see above).

The report concludes with an Appendix containing thirty-seven pages of tables. Most of these present comparisons of mean test scores and are rather confusing at first sight, since the two columns headed 'Sigma M1' and 'Sigma M2' are wrongly so described: these columns appear in fact to give the squares of the standard errors. The formula used for examining the significance of the differences between means is one which is strictly applicable only when the numbers in the two groups compared are large, say at least 50. This requirement is rarely satisfied and it would have been better to have examined significances by the *t* method, which would have made the probabilities given in the last columns of the tables rather less.

The report is based on a large amount of careful work and it is unfortunate that the presentation of it is so confused and confusing. The results obtained, where sufficient evidence to judge is given, seem to be similar to those of other workers.

E. G. C.

Developmental Abnormalities of the Eye. By IDA MANN; with a foreword by SIR JOHN PARSONS. Cambridge University Press. 1937. Pp. xi+444. 50*s*. net.

This is an extremely detailed and thorough treatment of all deformities in which the orbit is involved and of abnormalities affecting the eye alone. It is, in the words of its author, "intended...as an example of how a pure science may be linked with an empiric one (embryology with ophthalmology) to the advantages of the latter". It is not intended for the psychologist or the physiologist, and there is very little that will be of interest to him. Some examples, however, that may be cited of abnormalities that throw light on the functioning of the eye are congenital night blindness and Oguchi's disease, in which the retina exhibits a change of colour, accompanied by slowly increasing acuity, during darkness. This is attributed to pigment migration or to chemical changes in visual purple. Such a view is intriguing in view of the rôle suggested for pigment migration by Stiles and Crawford in accounting for some recently discovered phenomena. The sensitivity to light possessed by the muscles of the sphincter iridis is also of interest. K. J. W. C.

Les Yeux et la Vision. Par M.-L. Verrier. Paris: Librairie Félix Alcan. 1938. Pp. 171.

This is a most interesting study of the comparative anatomy of the eye, directed to showing that the traditional dualistic concept of rods and cones is unjustifiable and misleading. According to Mlle Verrier, the visual receptors are so manifold in nature that a polymorphism must be substituted for this dualism, and she hopes that it will lead to an improved outlook in the physiology of vision.

She advances arguments from behaviour and habitat as well as from anatomy, showing that there is little evidence of discontinuity and dual function, for instance, in the variation of visual acuity with illumination. She also criticizes the arguments based on the strictly nocturnal and diurnal habits of some species, pointing out that visual receptors are often very similar in members of a class, though some may be diurnal and others nocturnal, as in the case of lizards and vipers. Further, she considers that visual organs do not appear, in general, to be particularly well adapted to the habits and requirements of their owners, and that argument from mode of life to function of organs is thus illegitimate.

Mlle Verrier has written a most valuable and interesting account of the numerous types of visual receptor which may be met with; and there is independent evidence that the classification even of the receptors in the human eye is not so simple as it seemed; e.g. Lythgoe has concluded that peripheral cones differ functionally from central ones. There is, however, considerable evidence of two main types of receptor in the human eye—those in the fovea, which are capable of little dark adaptation, and those in the periphery, which are capable of a great deal, but have little colour sensitivity; and Steinhardt has found a suggestive inflexion in the curve showing differential brightness discrimination under varying illumination. Mlle Verrier considers the night-blindness found in vitamin-A deficiency to be merely a general effect of the malaise consequent upon avitaminosis. She is impressed by the fact that the rod-free area of the retina is sensitive to white light, though surely the simultaneous stimulation of complementary colour receptors can account for this. She holds that the decrease in acuity with falling illumination is mainly due to the bad definition afforded by the lens with a wide pupil; but even the dilation of atropine does not reduce resolving power to 30 min. of arc, as may occur under very weak illuminations. But these are minor points in a book which seems to be generally accurate and affords a wealth of interesting evidence. It is doubtful, however, whether comparative anatomy can give direct proof either way, or whether this is not more likely to result from physiological or psychological experiments.

K. J. W. C.

Recent Experiments in Psychology. By L. W. Crafts, T. C. Schneirla, E. E. Robinson and R. W. Gilbert. London: McGraw-Hill Publishing Co. 1938. Pp. xiv+417. 12*s.* 6*d.* net.

All of the authors, or compilers, of this book are on the staff of the New York University. They say that their aim is "to make available to the college student and to the general reader certain investigations which are representative of the experimental treatment of twenty-eight important psychological topics". The general reader will find the volume very heavy going, and since pretty well all the reports of experiments are re-written from perfectly easily accessible journals it is not at all easy to see why the 'college student' should not be expected to consult the originals where necessary. It is true that each experimental report is prefaced by an introduction which is said to be designed to orient the reader with respect to the topic dealt with. To many people the book will in fact seem like a bit of unnecessary spoon-feeding. The topics dealt with are extremely various, beginning with The Origin of the Cat's Responses to Rats and Mice and ending with The Judgment of Vocational Aptitude and Success from Photographs. It is not at all easy to detect any real thread of connexion from topic to topic or chapter to chapter. Nor is it easy to discern what principles, if any, led to the selection of the experimental reports presented. The great majority of the topics have, during the last few years, been made the subject of very many researches. It may be that one is meant to assume that the particular works selected are the best and are all of them of such permanent value that they ought to be enshrined in book form. That would be a very large assumption to make, even though it may be perfectly true that every research selected is good and deserves study. The fact is that current psychological research cannot really be properly represented in this manner. The general reader must have more simplified discussion and less technical detail; the student ought to go to the journals and be so instructed that he can find his own way about them.

Experimental Foundations of General Psychology. By WILLARD L. VALENTINE. New York: Farrar and Rinehart. 1938. Pp. xvii + 377. $2·00.

Prof. Willard Valentine addresses this volume to first-year students of psychology who desire to know something about the possible factual basis of many of the general statements which they may meet. On the whole the book should serve extremely well the purpose for which it is meant. The selection of topics is perhaps a little dominated by a bias towards the professedly practical and towards the *bizarre*, but there is no doubt that a great many students who come to psychology *are* interested in these directions, and if the interests are met fairly and without prejudice they may be made a basis for something more permanent and more serious. The great bulk of the experimental work described is work done by American investigators, but the descriptions are clear and lively. Very good summaries are added, and plenty of references to easily available literature.

The Kelley Statistical Tables. By TRUMAN LEE KELLEY. New York: The Macmillan Company. 1938. Pp. 136. $4·50.

Prof. Kelley has compiled a set of tables which should prove of considerable use to workers whose data permit of statistical analysis to a high degree of precision. The longest table gives Normal Distribution, Simple Correlation and Probability functions to eight places of decimals, and these are supplemented by cubic, quintic and septic interpolation coefficients, given to ten or eleven decimal places. The data of the laboratory or research psychologist usually do not warrant analysis to the high degree of accuracy which would justify the use of these tables, but there is much in the volume which would be very useful to anyone dealing with correlations and contingencies. The table of χ^2 functions, for example, is given in an unusual form which makes interpolation a relatively easy matter. For the full use of the tables an advanced standard of mathematical ability is required, but the ordinary research worker, guided by the explanatory matter, would find much of value to him in the analysis of his data. The book ends with tables of square roots to eight places and of constants frequently needed. Whilst these latter can be found elsewhere, of course, it is of service to have them bound with the other tables. The printing is extremely clear and well set out and the volume is of a convenient size and shape.

√ *Race: A Study in Modern Superstition.* By JACQUES BARZUN. London: Methuen and Co. Pp. x + 353. 10*s.* 6*d.* net.

This book is undeniably good. It might have been better still. The author, who writes with great force and humour and is both clear and concise, is concerned to show up the tremendous effects of centuries of 'race-thinking'. He gives most excellent accounts of a good many particular racial doctrines, of their impact on past history, and of their devastating appearance in current affairs. He has no difficulty in showing that all and any of these doctrines are riddled with error, unbalanced and insecurely based. He recognizes much more clearly than most writers that this makes little or no difference to their psychological significance and to their effects upon group policy and action. But he does not try to show seriously just how it is that the superstition of race gets so deep a hold upon popular thinking. Especially it seems a pity that he fails to consider whether there may be a sense in which very long-standing and well-organized groups acquire persistent characteristic qualities which go on from generation to generation, not in a genetic but in a psychological manner, and powerfully affect the relationships of these groups with others. If this is the case these qualities, though they are not racial, may, for example, reasonably be called 'national', and then there would be some basis other than the merely emotional one for many of the race doctrines. Still, even as it stands, the book is timely and should be widely read by psychologists and others. An appendix contains a beautiful collection of opinions called 'Race-thinking: a brief anthology', and there are some excellent bibliographical notes. The author is so obviously fitted for this kind of investigation that it is greatly to be hoped that he will follow up this admirable book with another one which will deal more particularly with the problems of the character qualities of groups, and whether these really do belong to the group organizations as such.

Emotion and the Educative Process. By DANIEL A. PRESCOTT. Washington: American Council of Education. 1938. Pp. xviii + 323. $1·50.

In 1934 an American Council of Education formed the project of carrying out an informed survey of the present state of psychological and educational knowledge and opinion concerning the character and basis of emotion and its relation to various practical problems as they are met in schools. This is a full report of the survey as it has been effected. After a statement of the general and particular aims of the undertaking the book discusses in order: Basic Affective Phenomena; Physiological Bases of Affective Experience and Behaviour; Patterning and Trainability of Affective Behaviour; Affective Maturity; Basic Personality Needs and Conditions which Frustrate Them; Affective Behaviour and Contemporary Social Institutions and Processes; The Influence of Affective Factors upon Learning; Affect and Education; Aspects of Education Needing Study; Personnel Problems in Education. It presents a useful bibliography. As befits its aim the book is concerned not with the organization of original research and reports upon this, but with collating and constructively criticizing a large bulk of work already carried out upon the topics selected for special consideration. It does this very well, on the whole stating the present position fairly concisely and criticizing it with force and justice. Thus any reader interested in the general subject can from this book get a pretty good idea about the issues of much of the recent work, at any rate of the recent American work, on the problems involved. Where the report is somewhat disappointing is in its formulation of programmes for additional research. Many proposals are made, and though practically all of these are good and very much to the point, they are made in so general a way that they will be of little real use to the prospective research worker. In this field, as in many others in psychology, it is less difficult to formulate problems, or to see where lie the problems that it is most urgently desirable to investigate, than to invent feasible and practical methods of research. It is, for example, not of much help to say that the whole concept of learning needs revision and "it is greatly to be hoped that this will be cleared up by the findings of penetrating research". The Committee has certainly done useful work in surveying the field and in showing where some of the main problems as yet inadequately solved are to be found. If now it would go on to suggest just how the methods of research that have been found most fruitful might be further developed, and to indicate new methods which could be applied with a reasonable hope of success, its work would be far more valuable still.

The Psychology of Early Growth. By ARNOLD GESELL and HELEN THOMPSON. New York: The Macmillan Company. 1938. Pp. ix+290. $4·00.

This is a guide book for the research worker who is interested in the controlled study of the growth of behaviour in young infants. It describes, illustrates and advises specific methods of investigation. It gives behaviour norms, maturity level summaries, a functional syllabus, and anthropometric norms. It presents principles of genetic analysis and shows how to achieve a clinical assessment of behaviour status. In appendices it gives many types of blank upon which the relevant material for this kind of study may be most conveniently entered. The whole provides a volume of intense interest and practical value to the research worker in this field, and it should certainly be in every psychological and educational library.

The Development of Linguistic Skill in Twins, Singletons with Siblings and Only Children from Age Five to Ten Years. By EDITH A. DAVIS. University of Minnesota Press, Institute of Child Welfare, Monograph No. 14. London: Humphrey Milford. 1937. Pp. x+165. 9*s.* net.

Four hundred and thirty-six children, coming from various American schools, were set in what are called standardized situations, and the remarks they made elaborately analysed for articulation, length, number of words and other characteristics. The situations comprised the presentation of a considerable number of toys which the children were encouraged to play with. All remarks made were recorded. Sometimes remarks had to be encouraged by the experimenter. It is claimed that only children are superior to others in linguistic skill, and that twins are especially backward, though the disability of twins depends to some extent upon economic status. Many early language inferiorities are really connected with imperfect articulation, and if this persists various other disabilities are likely to arise. The monograph is clear and well written and provided the method is accepted the analysis made could hardly be better than it is.

The Adolescent. By ADA HART ARLITT. London: McGraw-Hill Publishing Co. 1938. Pp. ix+242. 8*s.* 6*d.* net.

Dr Arlitt hopes that her book will help to stop parents from worrying unnecessarily about the vagaries of their adolescent children, and to understand and sympathize with them instead. Perhaps it will. Not a little of the terminology used, and many of the institutions and fashions referred to are American, but the English parent will probably find parallel illustrations which fit just as well without any difficulty. The general tone is sweeping and often dogmatic, and the parent who takes the whole thing very literally will have a lot to do, and will probably have to be well prepared to bear patiently or to rebut the charge of interference. It seems a pity that more specific evidence is not given for many of the statements made.

The Generations: A Study of the Cycle of Parents and Children. By EMANUEL MILLER. London: Faber and Faber. 1938. Pp. 276. 7*s.* 6*d.* net.

The topic of this book is the family. There is some study of its history, of its past and present functions and structure—mainly considered from a psychological point of view—of the individual demands which the family group has made and continues to make upon its members, and of its probable future. Moreover it is a good book which can be read by anybody who is sufficiently intelligent both with pleasure and with profit. It differs from a lot of books having a general interest in that it never descends to the level of offering cheap and popular advice. It is essentially a cool and careful study of the relevant facts, of methods available for their investigation, and of the main conclusions which the author considers may be regarded as established concerning them. Since the whole is based upon first-hand experience together with a thoroughgoing survey of authoritative investigations, it would have added to the value of the book if, when the latter are mentioned, the references to the literature had been given. But apart from this Dr Miller has produced a most readable, sensible and helpful book.

Psychotherapy. By PAUL SCHILDER. London: Kegan Paul, Trench, Trubner and Co., Ltd. 1938. Pp. 344. 15*s.* net.

This book is intended primarily to assist the physician and the psychiatrist in their approach to psychotherapy, though its description of psychotherapeutic method may be of considerable interest to the psychologist who does not intend to practise. The introductory material is difficult to understand, since it is linked to certain theories of the author's about the relation of the psyche to the 'body image'—theories developed in other books, but nowhere clearly stated here. However, what does emerge is the importance of considering the psychological symptoms connected with organic disease. There is a tendency to maintain, for instance, through the projection of physiological symptoms, the experience of health in conditions of severe organic disease. The nature of this tendency is related to the psychological make-up of the individual; and it is of great importance that it should be taken into consideration in diagnosis and treatment.

The main part of the book is devoted to a full and interesting description and discussion of the various techniques which the psychotherapist has at his command (discussion, free association, dream interpretation, etc.); and of the main psychotherapeutic systems and their application to different types of neurosis and psychosis. The author has developed an interesting method of group treatment, in which discussion, association and interpretation are carried out with as many as six or seven patients together. The method is said to be successful even with severe neuroses.

The book is suggestive and interesting, especially in that the author is not wedded exclusively to any psychoanalytic school. But his style of writing is disjointed, and his method of presentation, especially of the various types of neurosis, etc., disconnected and unmethodical; so that it is often a matter of difficulty to follow his line of thought.

Personality in the Depression. By E. A. RUNDQUIST and R. F. SLETTO. University of Minnesota Press; London: Humphrey Milford. 1936. Pp. xxii + 398. 18*s.* net.

Both the personality and the depression referred to are American. A large number of employed and unemployed persons of both sexes filled up personal forms and recorded their views about a variety of topics on the Minnesota Scale for a Survey of

Opinions. A number of statements such as Times are Getting Better, or It is Difficult to Say the Right Thing at the Right Time, are printed. The patient is asked to record on paper, by underlining, whether he strongly agrees, agrees, is undecided, disagrees, or strongly disagrees. There are also a Moral Scale, an Inferiority Scale, a Family Scale, a Law Scale, an Economic Conservatism Scale, and an Education Scale. On the whole the unemployed differed from the rest most in their reactions (paper) to the economic order, and discontent with this appeared greater among men of the more lowly occupations, the least education, and among the older subjects. But probably the main value of this publication consists in its critical treatment of the methods used. These methods are given a thorough trial and are well analysed. They are found wanting in a number of specified respects, and as a basis for further research on *method* the monograph will be found useful.

L'Étude Expérimentale du Caractère: Méthodes et Résultats. Par N. BRAUNSHAUSEN. Vossegat: Centre National d'Éducation. 1937. Pp. 198.

The English reader who wishes to get a brief but clear account of many of the recent developments of characterology will find this book exceedingly useful. It gives first a general account of the development of experiments on willing, with sufficient detail to show up their methods and their main conclusions. There follows a really excellent description of the various typologies which have grown up in post-war days, especially in Germany. Then comes a study of the methods used in the investigation of more or less isolated character qualities, followed by a chapter dealing with the ways in which the study of somatic reactions may help to elucidate character and temperament. Finally various views, partly experimental, partly observational, on the general basis of character and personality and the interrelationships of their different aspects are set forth. Every chapter has a brief and clear summary, and in the end the whole position is carefully summed up. The author does not attempt to give an adequate account of work that has appeared in the English language, but his study of French, Belgian and German developments is very welcome. The bibliography also should be of great use to the English student. Throughout, a very sane attitude of appreciative criticism is maintained.

Charakterologie. Von PAUL HELWIG. Leipzig: B. G. Teubner. 1936. S. xii+295. geb. RM. 8·60.

The English reader who desires a fairly brief and comprehensive study of modern German developments in what is called Characterology will be able to satisfy himself with this book. After rather a lot of general remarks about the aims and methods of the study of personal character, the author gives an excellent survey of most of the main typologies, continues with some account of psychopathology, summarizes the work of Klages, and considers the contributions made by psychoanalysis. He also discusses the more philosophical approach of Haberlins, and some of the by-products of the movement. There is exceedingly little sign of any influence coming from work done outside Germany.

Philosophie und Pädagogik Paul Häberlins in ihren Wandlungen. Von PETER KAMM. Munchen: Ernst Reinhardt. 1938. S. 479. M. 6·80.

Haberlins' philosophical and educational views are here expounded in a very appreciative manner by Peter Kamm. They are considered in the order of their development and the author sets them in a framework of four periods: (1) the period during which the basic religious concepts which underlay all the later developments were forming (1907–8); (2) the period during which Haberlins was more immediately concerned with the idealistic and moral implications of his ideas (1910–25); (3) the period of the full expression of his religious philosophy (1926–32); and (4) the period in which theoretical philosophy formed the main pre-occupation (from 1932). Throughout stress is laid upon the educational significance of the views expounded, and the final section of the exposition deals explicitly with the pedagogical writings of 1935–6. A portrait of Haberlins is included and a full bibliography of all his published works.

Science and Psychic Phenomena. By G. N. M. TYRRELL. London: Methuen and Co. 1938. Pp. xv+379. 12*s.* 6*d.* net.

The student who desires to find in one book a fairly detailed study of most of the types of problem which have been investigated in the name of psychical research could hardly do better than consult this volume. Mr Tyrrell considers, with abundant illustration, the case for different kinds of 'spontaneous evidence', both outside and inside the laboratory; he discusses possible physical explanations of the alleged facts and cautiously rejects them; he passes in review the evidence concerning many different kinds of trance phenomena. While in the end he says frankly, "I do not think that we can point to any fact in psychical research to-day and say that it is universally accepted, unless, indeed, we regard hypnotism as being a psychical phenomenon", he considers himself that the evidence for extra-sensory perception, incapable of explanation in physical terms, is practically convincing, and that the case for the "existence and communication of the surviving self-principles of deceased persons is strong". He suggests hypothetical explanations.

Psychology and Religion. By C. G. JUNG. New Haven: Yale University Press; London: Humphrey Milford. 1938. Pp. 131. 9*s.* net.

This book is based upon a series of lectures delivered at Yale University on the Terry Foundation. Prof. Jung is concerned to relate together religion, defined as the experience of the 'numinous', and the functioning of the collective unconscious. The latter is illustrated in particular from a long series of dreams of a certain neurotic patient, which contained symbolisms characteristic of the occult religions, such as Gnosticism—symbolisms quite unknown to the patient's conscious mind. These culminated in the visual impression of a 'world clock', producing in the patient "an impression of the most sublime harmony". From this and from many similar 'mandalas' produced by patients, Prof. Jung concludes that the religion of man is no longer based on the projection of the 'numinosum' upon an external and autonomous deity, but upon the idea of the complete man; "the unconscious produces a new idea of man deified or divine, impersonal, concealed, protected, usually dehumanized and expressed by abstract symbolism".

It is impossible to give in such a brief account any idea of the subtlety and profundity of Prof. Jung's argument, which requires close and detailed study. It will be unfortunate if the excessive price of this book prevents many who would otherwise do so from studying it for themselves.

Die geistige Wiederholung. Von F. A. VAN SCHELTEMA. Leipzig: Bibliographisches Institut AG. 1937. S. viii+294+xxxii. M. 5·80.

Dr Van Scheltema's new book may be called a study in comparative anthropology, or, using the words in the old sense, in genetical psychology. Its general object is to trace the ways in which the same motives, technique and material appear again and again in cultural periods which are widely separated in space and time. Most of its pages are devoted to a study of artistic expressions as revealed in the works of ancient times and that of contemporary child drawings. But a good many other departments of culture are touched upon as well. A general parallel is drawn between culture development and individual development, and attempts made to state in what respects the two run along similar lines and in what respects they differ. The volume is well illustrated by 32 pages of plates.

Theodor Elsenhans Lehrbuch der Psychologie. Herausgegeben von HANS W. GRUHLE und F. DORSCH. Tubingen. 1938. S. 241–400. RM. 6·30.

Elsenhans's text-book, fully revised from the third edition by Prof. Fritz Giese, is being issued in parts by J. C. B. Mohr (Paul Siebeck) of Tubingen. This particular part deals, in a rather general way, and with critical and expository references, with many of the main theories of perception, thinking and feeling.

Deuxième Congrès International d'Esthétique et de Science de l'Art, Paris, 1937. T. I et II. Paris: Félix Alcan. 1937. Pp. lxiii + 369; 527. 120 f.

General introductory and laudatory orations were delivered by MM. Paul Valéry, Paul Claudel and Victor Basch and these are here printed in full. Fairly extensive summaries are given of the numerous other contributions, and these are arranged under the following headings: General Aesthetic; Psychology, analytic (i.e. more or less experimental) and general; Sociology and Culture; History and Criticism; the Science and Technique of Art, and Contemporary Art. Every reader who is interested in problems of artistic production and appreciation will find something to attract and intrigue him. Work from twenty-three different nations is represented. The volumes are published under the general direction of the Association pour l'Étude des Arts. Most of the psychological contributions are concerned with some problem or other closely related to the emotional life, and a good many of them show some leaning towards psychopathological theories. As would be expected, they display a considerable diversity of opinion and conclusion, and in these ways they are undoubtedly stimulating and can be utilized to suggest many lines of work.

La Philosophie de Heymans. Par T. J. C. GERRITSEN. Paris: Félix Alcan. 1938. Pp. 299. 50 f.

Many English students must have found their way to the University of Gröningen at some time or another during the long period when Heymans held there the Chair of Philosophy and Psychology. To these, all of whom probably remember a kindly reception and stimulating companionship, this book will be welcome, and doubtless to many others as well. It deals, indeed, only with Heymans' philosophy, and that is a pity, for his psychological work, which was attractive and influential, is still less well known than it deserves to be. M. Gerritsen deals in turn with Heymans' theory of knowledge, his ethics, his aesthetics and his metaphysic, and in each case his expositions are clear and appreciative. The book concludes with a bibliography of Heymans' published writings.

The Pendulum Swings Back. By MARVIN M. BLACK. Nashville: Cokesbury. 1938. Pp. 229. $2·00.

The 'swing back' to which the title of this book refers is a recoil from mechanism. In his foreword Prof. Ellwood says that the author presents "a synoptic view of the results of all the great sciences which are antecedent to the sciences of man and of human society". Actually there is not very much about these antecedent sciences, for the bulk of the book is directly concerned with psychology, psychic research, medicine, social philosophy, journalistic education and cultural anthropology. In so far as biological problems are discussed, they are for the most part closely related to human affairs. The book may be useful as a kind of summary of what non-materially minded investigators and thinkers are saying about human problems, especially in regard to their social significance. But it is not very easy or attractive to read. There are far too many extremely summary references to most varied authorities, from which no reader who does not know already a good deal about their work will be able to gather very much. The general tone also is somewhat dogmatic.

A Biological Approach to the Problem of Abnormal Behaviour. By MILTON HARRINGTON. Lancaster, Pa: The Science Press. 1938. Pp. 459.

Dr Harrington believes that "the psychopathology of the future must be mechanistic not motivistic", and that everything must be built upon a physiological foundation. His basic tenet, he says, is that behaviour "is due to the action of an anatomical structure or mechanism of some sort, and that abnormal thoughts, feelings and actions are due to the limitations and defects of this piece of mechanism". Part I discusses First Principles and does very little more than assert this position over again in very general terms. Part II is called Psycho-Physiology. A large number of different topics are dealt with here, from a consideration of very simple forms of behaviour to that of very complex intellectual, emotional and sentiment reactions. The effort all the time is to sketch a general physiology of the neural and other conditions upon which each form of conduct that is being dealt with may be regarded as based. The assumption is that the assigned physiological basis is also a veritable cause. Much of the discussion will be found interesting and helpful by anybody who has a sufficiently good physiological background to read critically. But of necessity many of the 'basic conditions' considered are described only in an extremely vague and general manner. Again and again very little is done except to translate what is known directly in terms of the description of gross observed behaviour into terms of what is suspected from the physiological study of reactions vastly simpler than those which are being considered. Part III concerns Psychopathology. Here once more the discussions are of a general kind concerning *principles* of defective functioning rather than specific cases.

Most readers will probably feel that as a scheme for an approach to the study of problems of behaviour this may be all very well. They will find also that if they try to express all their particular problems in specific terminology consistent with the scheme they will often be baffled and have to resort to much speculation. As an attempt to develop an orientation and to guide research the book has much to be said in its favour. As a statement of anything in the least final in the way of theorizing it is unsatisfactory. Moreover it is much too long.

Identité du Monde et de la Connaissance. Par GEORGES MATISSE. Paris: Félix Alcan. 1937. Pp. 256. 35 f.

A tremendous number of problems are attacked in this book, but always with vigour and with grace. In fact very nearly the whole mechanism of the mental life, from sensations, both exteroceptive and proprioceptive, through affect and emotion, to perception in all forms and the higher developments of the reasoning processes, is passed in review. Also the author has much to say about aesthetic reactions and about many kinds of mysticism, especially those which have been connected with religious beliefs. In general his aim is to decide which of all these processes give reasonable and more or less well attested knowledge of nature, and which represent rather by-paths of exploration and the pursuit of extremely attractive but romantic problems. The volume will admirably serve the purposes of the reader who wants to be stimulated, amused, interested; as a serious scientific contribution to psychology it is not satisfactory. But then it does not pretend to be this.

L'Alternative. Par VLADIMIR JANKÉLÉVITCH. Paris: Félix Alcan. 1938. Pp. 219. 30 f.

All life and every moment of it is Alternative. With plenty of decorations and variations this theme is played throughout the first half of the book. There is no escaping alternatives in their many forms, but it is very interesting to know something of how they arise, and in general they can be borne, and even mastered, if only the patient man is sincere. The second part of the book deals with the metaphysic of boredom. Boredom also is pretty nearly a universal lot. The only thing to do about it is to face it squarely and to learn that every minute has its peculiar worth and excellence. "Les hasards charmants naissent et se fanent dans toutes les allées, il ne faut pas bâiller pendant ce temps-là."

Social Interest: A Challenge to Mankind. By ALFRED ADLER. London: Faber and Faber. 1938. Pp. 313. 10*s.* 6*d.* net.

A number of related essays or lectures on a variety of topics by the late Dr Alfred Adler are here presented to the public. By the person already familiar with Dr Adler's work the book will not be found to contain any new ideas, but it gives many excellent, clear and concise summaries of his views, and some new case material.

The Voice of Experience. London: Herbert Jenkins, Ltd. 1938. Pp. 203. 5*s.* net.

The publishers modestly remark that "The Voice of Experience" is a "psychologist whose knowledge of human nature is unequalled throughout the world". This Voice here purports to tell everybody how to be extremely happy. The book is costly.

VOLUME XXIX JANUARY, 1939 PART 3

CALCUTTA UNIVERSITY CENTRAL LIBRARY

OBITUARY NOTICE

PROFESSOR WILHELM STERN, 1871–1938

WE regret to announce the death of Professor Wilhelm Stern, an Honorary Member of the British Psychological Society. He had been associated mainly with the University of Hamburg. But he was forced to leave Germany, and went to work in the Department of Psychology at Duke University. He is best known in this country for his book, *The Psychology of Early Childhood* (first German edition, 1914), which is regarded as a standard work on the subject. His earlier work on applied psychology—mental measurement, vocational guidance, etc.—was succeeded by his main life's work on the 'science of personalistics'. This included approaches from the psychological, physiological, cultural and metaphysical aspects of 'personalism', which are embodied in a three-volume work, published between 1906 and 1924, *Person und Sache: System des kritischen Personalismus.* It was further developed as a systematic psychology in *Allgemeine Psychologie auf personalisticher Grundlage* (1935), translated and published in 1938 as *General Psychology from the Personalistic Standpoint.*

THE CO-OPERATION OF SOCIAL GROUPS. PART I

By D. M. CARMICHAEL

From the Cambridge Psychological Laboratory

I. *A scheme of research* (pp. 206–209).
  (i) *Direct observation of limited communities* (pp. 207–208).
  (ii) *Controlled experimentation* (p. 208).
  (iii) *Experiments in constructive social thinking* (pp. 208–209).
II. *Further experiments in constructive social thinking* (pp. 210–215).
  (i) *Material of Case I: "Church Union in Thorness"* (pp. 210–214).
  (ii) *Method of presentation* (p. 215).
III. *Analysis of results of Case I* (pp. 216–231).
  (1) *Numerical analysis* (pp. 216–218).
  (2) *Qualitative analysis* (pp. 218–231).
    (i) *Initial union* (pp. 219–222).
    (ii) *Permanent union* (pp. 223–225).
    (iii) *Impermanent union* (pp. 225–226).
    (iv) *Harmonious union* (pp. 226–227).
    (v) *Union with disharmony* (pp. 227–228).
    (vi) *The fisher-folk* (pp. 228–229).
    (vii) *The influence of a leader* (pp. 229–231).

I. A SCHEME OF RESEARCH

EXACTLY 200 years ago David Hume conceived of a science of man based not on unverifiable metaphysical speculation but on accurate observation and the use of the experimental method. In this science "we must glean up our experiments from a cautious observation of human life, and take them as they appear in the common course of the world, by men's behaviour in company, in affairs and in their pleasures. When experiments of this kind are judiciously collected and compared, we may hope to establish on them a science which will not be inferior in certainty, and will be much superior in utility, to any other of human comprehension."

Hume's high hopes have scarcely been realized. General psychology, it is true, has gone a long way towards perfecting experimental methods even more objective than those which Hume seems to have in mind, but

CALCUTTA UNIVERSITY CENTRAL LIBRARY.

of the utility of much of this work the ordinary man in the street has still to be convinced. In social psychology, on the other hand, where the utility of scientific investigation is more immediately apparent, the complexity of the data and the difficulty of performing controlled experiments are such that individual interpretation and speculation still hold the field, and "even the rabble without doors may judge from the noise and clamour which they hear, that all goes not well within".

The experiments described in this paper form part of a wider scheme of research devised and inaugurated at Cambridge by Prof. F. C. Bartlett. This scheme aims ultimately at discovering, as far as possible by experimental methods, the psychological factors which govern social harmony and discord, with special reference to those conditions which make for permanent and friendly co-operation between differently organized social groups. A solution of this all-important question is being sought at least as much for its practical value as for its academic interest. The methods of approach which it is proposed to use, and the reasons for their adoption, can best be indicated by giving a brief summary of a "Preliminary Report and Suggestions" which Prof. Bartlett has already published.[1]

Three interdependent lines of research are suggested, of which the third is intended to give direction to the first and second, and the first and second to validate the third.

(i) *Direct observation of limited communities*

"There is a large number of limited communities within which the friendly co-operation of different social groups is demanded", either permanently or from time to time. The history, outstanding membership and cultural background of many of these small groups is either already well known, or can easily be determined, and the interaction of the groups can therefore be at once made the subject of first-hand observation. This observation will be directed towards discovering "what are the most common *human* conditions which either maintain, or hinder, friendly co-operation".

All such observational work will, of course, "have to pay attention to certain principles of method in the collection and recording of facts", and it would be desirable to have the facts "available at some common centre in a form which renders their evaluation possible". This, it is interesting to note, is precisely the method proposed by Hume—the

---

[1] "The co-operation of social groups" (1938), *Occupational Psychol.* XII, 30–42.

"judicious collection" of relevant data "as they appear in the common course of the world", and their subsequent comparison and utilization.

(ii). *Controlled experimentation*

Within the type of community mentioned, it is possible to stage social experiments either with groups which are already in contact, or with groups which have been brought together 'for the occasion', and encouraged to co-operate[1]. In both these instances, the experimenter could "exaggerate, or otherwise alter certain outstanding cultural or individual conditions which are suspected to be of influence, and then keep a watch upon their operation".

As it seems to be a reasonable and necessary working hypothesis to suppose that the *principles* which govern the co-operation and conflict of groups remain much the same, whatever the size of the group, the fact that observational and experimental work can be carried out only with small groups is relatively unimportant. What is important, however, is that this observational and experimental work should be properly directed. The mere amassing of data defeats its own ends, and the conducting of significant experiments becomes impossible unless some principle of selection is used which enables us, within limits, to distinguish between those factors which are relevant and those which can safely be neglected. The third method of approach, now to be described, will, it is hoped, provide principles of selection.

(iii) *Experiments in constructive social thinking*

This method aims primarily at discovering what factors are *commonly believed* to be of leading importance in the relationships of groups. In a sense which will become apparent, it is essentially an attempt to discover "how people think about social questions". Bartlett (*loc. cit.*) describes it by stating the manner of its adoption: his experiments on remembering having revealed "certain very important processes of mental construction", he thought that "it would be interesting to get to closer grips with these processes, to find how and when they occur, to distinguish between the part of them which is contributed by the individual and the part, if any, which has a universal determination". With this end in view, he "devised and applied a considerable variety of stimulus material", including certain (written) descriptions "in which situations were depicted involving more or less indeterminate relationships between two or more people and the subjects were asked to carry

---

[1] Cf. Lewin, K. and Lippit, R. (1938). *Sociometry*, I, 292–300.

CALCUTTA UNIVERSITY CENTRAL LIBRARY.

further the probable issue of these relationships". It was noticed that "when they did this they very often made use of facts of a genuinely cultural or group character, which could not be expressed in terms of individual behaviour or experience alone". It was therefore decided "to have instances involving direct group relationships, and to get the subjects to think constructively about them".

A number of concrete situations were therefore described, all involving group relationships not unlikely to occur within the experience of the subjects available. These situations "were presented in an incomplete form, and the subjects were asked to carry further their probable history". The "Preliminary Report and Suggestions" prints two of the descriptions used by Bartlett, and contains an analysis of the results which they yielded. It is unnecessary here to discuss these results in detail, as one of the descriptions in question has since been used more extensively, and reference to it will be made at a later stage in this paper. But Bartlett's general conclusions as regards his new method are important. They are three in number:

(1) "When an ordinary person is confronted by a concrete description of a social situation in which agreement between two groups is sought, but some discordant elements are present, he can usually give a considered opinion as to the probable upshot." The method has, for this reason, certain marked advantages over most questionnaires.

(2) "The method shows that there can be a considerable measure of agreement regarding the primary bases of social harmony and discord, and at the same time a wide diversity of view about the particular mechanism through which these fundamental conditions find expression." A further and wider application of the method will, it is hoped, sharpen and define this important practical distinction.

(3) A fairly circumscribed application of the method has already demonstrated that "ideas about the leading conditions governing co-operation and discord may vary significantly from one section to another of what is considered the same general community".

Widespread views about social co-operation must undoubtedly affect the chances of social co-operation, and therefore the opinions elicited by our third method will be to that extent of practical value. Whether the opinions are true or false, however, can be determined only by the use of the first two methods of direct observation and controlled experimentation. If, by the combined use of the three methods we are ever able to discover the factors which actually govern social co-operation, then the practical value of the whole enquiry will be immensely increased.

II. FURTHER EXPERIMENTS IN CONSTRUCTIVE SOCIAL THINKING

I shall now proceed to describe and discuss further experiments which have been carried out along the lines suggested by Bartlett. Three 'descriptions' will be referred to—one which Bartlett has already tested; one, a leading article from *The Times*, dealing with an actual attempt to secure co-operation between two sporting associations; and one of my own composition, which is given below as 'Case I'. In each instance I shall (1) reprint the material used; (2) state the reasons for its adoption and the manner of its presentation; (3) give a brief analysis of results and attempt to assess their significance. Finally, I shall endeavour to estimate the value of the work which has so far been done in relation to the wider scheme of research of which it forms a part.

(i) *Material of Case I: "Church Union in Thorness"*

In 1929, after a long series of legal, financial and administrative adjustments, and one or two minor changes in dogma, the United Free and the old Established Church of Scotland became in name and in the eyes of the state, one church. The day was indeed a memorable one when, for the first time, the representatives of each Church met as one body, and the first Moderator of this first General Assembly of the new 'Church of Scotland' rose to address its members. He spoke of the work which had been done, and of the success with which it had been crowned; of the difficulties which had been overcome; of the unselfishness and ceaseless vigilance of those to whom had been entrusted the task of framing the Act of Union; of the inspiration and guidance which had been granted them as they strove to realize their high ideal, and of the whole-hearted support which they had received from ministers of both denominations, and from eminent laymen within the churches. It was impossible, he said, to consider these things, and to look upon that vast assemblage, and not to feel convinced that, within the body of the ministry at least, the old prejudices, the old hatreds, the old jealousies had perished, and along with them all the waste, the bitterness, the pride and the self-righteousness born of two rival traditions. It was unnecessary for him to remind them that these traditions had grown out of a common stem: that this separation of the churches, now happily reunited, had lasted scarcely 100 years; that recently, as ministers of the gospel, they had tended more and more to forget their differences and to unite in one common purpose. And now, richer in experience and strengthened in faith, they were offered, by the grace of God, a golden opportunity to build up a united church worthy of its name—the Church of Scotland.

"Much, then, has been done," he continued, "and we have good reason to be thankful. And yet in truth our work has scarcely begun. For this Union of ours will be but a sham, a hollow mockery, unless in each and every parish in the land, congregations which hitherto have eyed each other with distrust and suspicion, seek a new, a deeper understanding—forgiving insult and injury, living down hard memories, overthrowing old idols, eradicating those sentiments which lead to dissension and disintegration, and so learning, slowly but surely, to work together in harmony for the advancement of the Kingdom. And although no two congregations will ever be forced to unite, yet, when opportunity offers, and union seems advisable, the congregations concerned will be asked to unite under one minister who has proved himself acceptable to both. It is those congregations above all others which will have to make sacrifices, to exercise real self-denial, to subordinate their parochial interests to the welfare of the church as a whole, and in general to show true love and forbearance. But theirs will be the greater reward."

Some years later, in Thorness, a town of some 2500 inhabitants on the north-east seaboard of Scotland, the difficulties of putting this ideal into practice became abundantly clear. In this instance, the difficulties had, in the main, two causes—the peculiar temper of the people, and a most remarkable combination of circumstances.

The temper of the people is best indicated by a description of their environment, for they seem to have taken on many of its qualities. To a stranger, Thorness appears as a cold, hard, business-like place, swept clean by the winds off the sea. Most of the houses are built of solid, grey sandstone, and the newer part of the town is well-planned, the streets being broad, well-paved, and set at right angles to one another. The old town, down by the river, where the fisher-folk dwell, is more ramshackle and much less prosperous-looking. These fisher-folk comprise less than one-quarter of the total population. Most of the men are fine, hardy fellows and expert seamen, and the women alert and active, but of late, owing to the failure of the fishing industry, they have fallen on evil days, and poverty and enforced idleness has brought about some degeneration in character. They are never great 'church-goers', although, like fisher-folk elsewhere, they are prone to short-lived bursts of intense religious enthusiasm.

The new town has grown up as a market centre for the rich agricultural district which surrounds it, and its population therefore consists mainly of artisans and middlemen, with a sprinkling of bankers, lawyers, doctors, school-teachers and retired farmers. It is these middle-class people who control church affairs. They are hospitable enough in their way, but just rather than generous, upright rather than amiable; direct in speech as in business; self-reliant, undemonstrative, very 'dour' if crossed. In religion they are moralists rather than mystics, showing little appreciation of, or reverence for, institutionalism or ritual as such. And, as one would expect, they are no respecters of persons. Perhaps they inherit their characteristics from their Norse ancestors, who founded Thorness 1000 years ago, and intermarried with the native Celtic stock. It is at least certain that the early descendants of this union, when oppressed by a worldly-minded bishop, remarked tersely, "Short rede is good rede; slay we the Bishop", and proceeded to roast him in the greater barn which he had built to hold his tithes.

There are seven churches altogether in Thorness, not counting the Salvation Army, the Faith Mission, and a few disconsolate Plymouth Brethren left high and dry by the last tide of revivalism. When the Act of Union was passed, the churches whom it did not concern—the Episcopalian, the Congregational, the 'Wee Frees', and the Original Seceders—waited with a certain cynical interest to see how it would affect the conduct of the other three. They changed their names, of course, the old Established Church becoming St Ninian's Church of Scotland: the first United Free becoming St John's Church of Scotland: and the second U.F. becoming St David's Church of

Scotland. But for a couple of years nothing exciting happened. Each church retained its own minister, St Ninian's remained, as always, somewhat aloof, and one could feel, beneath an outward show of friendliness, the same old undercurrent of jealousy between St John's and St David's. Of the two, the former had always been richer and more vigorous, and there was never any question of its being united with another church. But when the minister of St Ninian's—he was an old man—became weak and ill, and his congregation began to dwindle, as his preaching lost its power, the Presbytery decided that, on his retirement, every effort should be made to unite St Ninian's and St David's. The Convener of the Committee on Union pointed out that St Ninian's would be invigorated spiritually; St David's (and the Church of Scotland as a whole) would benefit economically. St Ninian's possessed one of the finest church buildings in the north of Scotland, well capable of housing both congregations; the church of St David's was small, and falling into disrepair, and those who worshipped there were too poor to restore it. It was true, of course, that neither church had ever shown an active neighbourly interest in the other. Yet at least they did not openly dislike one another, and the minister of St David's had expressed his willingness to be transferred, if necessary, to another charge, so that all the old associations might be broken off, and his congregation given every chance to develop a real unity in loyalty to a new minister. The cynics whispered that this was just another attempt on the part of the former U.F. ministers—they had a large majority in the united Presbytery—to bolster up one of their weaker congregations by allowing it to fasten on the wealth of the older church, and they predicted that the new minister would be an 'ex-U.F.' man, rescued from some hungry parish in the Highlands. But in spite of its detractors, the scheme would probably have gone through smoothly enough—perhaps a little lackadaisically—if it had not been for the extraordinary series of events already mentioned.

What happened was this. At the request of the minister of St Ninian's, who was due to retire in six months' time, a young man, MacKeen by name, was sent north from Edinburgh to help him in his preaching and in his parochial work generally. The appointment was understood to be a temporary one, so no very exact inquiry was made into the new assistant's credentials. His first sermon made a tremendous impression on a meagre handful of loyal worshippers; his second, preached in the evening of the same day, was, as the local papers put it, "much appreciated by a well-filled church". On the following Sunday those who came late had to stand, for all seats were taken. His width of appeal was amazing. Not only did staid middle-class Thornessians wax eloquent in his praise, but the fisher-folk from 'down the town' came over and over again to hear him, and people from the surrounding countryside who had no connexion whatsoever with the church could be seen hurrying townwards every Sunday morning—in cars, on bicycles, in 'buggies', on horseback and on foot. Not only could Mr MacKeen preach popular sermons, he could also sing. Every Sunday evening his congregation was "favoured with a sacred solo from the pulpit"; and seldom did a week pass but he penetrated to the remote fastnesses in the interior of the country, there to win encore after encore on some rustic concert platform. Some of the shrewd old elders of St Ninian's were none too sure of their man, distrusting especially his love of publicity, but they held their peace. The end of it was that the revitalized congregation petitioned the Presbytery that St Ninian's should be allowed to retain its independence (there being literally no room for another congregation) and that Mr MacKeen should be appointed assistant *and successor* to the old minister.

Whereupon the Presbytery had to take the matter seriously in hand. They found that owing to an oversight on the part of the authorities in Edinburgh, and a lamentable slackness in their own procedure, they had let loose in St Ninian's a man who, by the laws of his church, was not even qualified to be an assistant, much less to be ordained as a minister in full charge of a parish. They found other things also—a story of failure as a student and brilliance as a preacher; of numerous departures from one part of the country to another; of a short-lived blaze of popularity in each district and a willingness to be employed by a church of almost any denomination. Partly out of a desire to spare MacKeen, and partly in order themselves to avoid dangerous publicity, they rested their case on the technical point alone, and 'with much regret' asked MacKeen to resign.

Practically the whole town of Thorness rose up in arms against this verdict. "We have no time", they said, "for ministers of the gospel who hinder the preaching of the gospel." Beyond this they could not, or would not, see, except to speculate as to the real motives behind the decision, for they rightly suspected that so much good work would not be cut short, and so much ill-feeling aroused, for the sake of a mere technicality. The more outspoken of them had much to say of meanness, jealousy and self-interest, and the noise of the dispute was heard far beyond the borders of the country. Newspaper men hurried to Thorness; Mr MacKeen interviewed them all: and the Sunday papers revelled in this "Terrific Religious Crisis in the Far-flung North."

Its climax was indeed dramatic. On the Sunday before that on which he was due to resign, Mr MacKeen announced his intention of preaching a farewell sermon on the following Sunday. The Presbytery asked him not to do so, and they requested a minister from a neighbouring parish—one McTavish by name, an oldish man, who'd been a bit of an athlete in his day—to appear in the pulpit of St Ninian's. On that famous morning a sight was seen which will never be forgotten by those who witnessed it—two ministers racing for the same pulpit, and the older man getting there first by a good five yards! Mr MacKeen took his place in the congregation. The church was, of course, full to overflowing, and a curious crowd was straining round the doors. Mr McTavish rose to announce the first psalm. Before he had opened his mouth, Mr MacKeen rose also, and signed to his congregation to do the same. Then, with a look of Christian resignation on his face, he led the way out. Some of his rowdier followers as they went began to curse the Presbytery, the church and all that it contained, but Mr MacKeen rebuked them, and they filed out in silence. When the last of them had gone, Mr McTavish's voice echoed emptily over the heads of a dozen worshippers—the shrewd old elders, and three or four old ladies who were too rheumaticky to move.

After that the Presbytery could afford to let things take their course. They saw MacKeen hire a hall, and set himself up as an independent preacher. They saw him retain his popularity for another six months. Then little by little he began to lose his support. The zest of battle was gone, and the feeling of being martyrs in a common cause soon loses its binding force when one's self-denial gets no publicity. So first of all the fisher-folk lapsed into their usual Sabbath lethargy, and the country people lost their enthusiasm for Sunday travel. The poorer families found it very expensive to have to support a church of their own, and both they and the wealthier citizens among MacKeen's supporters could scarcely close their ears to the ever-increasing stock of rumours concerning their minister.

The Presbytery waited nearly a year. By that time most of those who had once been staunch members of St Ninian's had deserted MacKeen, and were, so to speak, sitting on the spiritual fence. The Presbytery succeeded in finding a young man (trained in the now "*United* Theological College") who was likely to make a sound and lasting appeal to any decent congregation. He was asked to stand as candidate for the united church of St David's and St Ninian's, the former minister of St David's having been, at his own request, transferred to another parish. St David's accepted him gladly: what was left of St Ninian's did so also, for any state was better than their present one. The formalities of union were completed, and the congregation of St David's entered boldly into the almost empty church of St Ninian's. The new minister, on being introduced to his united congregation, referred tactfully but firmly to the recent regrettable dispute between laity and regular clergy. No blame was apportioned on either side, but he hoped that, out of apparent evil, good would come, and that out of disruption a deeper and more lasting unity would be born. Perhaps in the course of time he would win back to St Ninian's not only those right-minded people who had been so strangely misled, but also some of those who had found their way there temporarily, and were now churchless once again.

Already, three months after his appointment, he has gone a long way towards realizing his ideal. The 'right-minded people' who had been sitting on the fence slipped easily into the fold, where they had the pleasure of reclaiming the best seats, and of reasserting their place in church politics. Then MacKeen, for reasons into which we need not enquire, had to leave Thorness; but even before he went many of his supporters had gone over to St Ninian's, preferring the solidarity of organized religion to the instability of an isolated church. A keen yachtsman himself, the new minister handles his fisher-folk as skilfully as he handles a boat, and they are now developing an interest in religion so well-grounded that it seems likely to be permanent. And in fact good has come out of evil, for the recent turmoil seems to have purged the town of strong emotion, the congregations of St David's and St Ninian's have so far fused with no wild outbursts of feeling, and little or none of the bickering and petty jealousies which so often make true co-operation impossible.

This is the longest and most complex example so far used. In writing it, I endeavoured to give a narrative of 'plain fact', any 'explanation' of the sequence of events being thrown in as a suggestion rather than as a true interpretation. The somewhat high-sounding 'Moderator's speech' is added for two reasons—to give historical perspective and (more important) to place in juxtaposition the real and the ideal, and to see whether in explaining what had occurred, and in predicting what was likely to occur, any of the selected groups of subjects would stress the high motives set forth in the speech.

(ii) *Method of presentation*

This example has so far been presented to 100 subjects. These fall into five groups:

English:	*Academic:*		25 University students, research workers and lecturers.
	Non-academic:		15 'W.E.A.' pupils, mostly engaged in practical, wage-earning activities.
Scottish:	*Academic:*		20 University students (75 % studying psychology).
	Non-academic:	*Clerical:*	20 ministers of the Church of Scotland in various parts of the country.
		Lay:	20 middle-class Scots, nearly all 'church-goers'.

The average age and age distribution in the English groups are closely similar; in the Scottish groups, the average age of the clerical group is rather higher than that of the lay, and much higher than that of the academic.

With the exception of about 50 % of the English academic, and 75 % of the Scottish academic, all replies were obtained by correspondence. The English exceptions dealt with the material in individual interviews, and the Scottish as a group.

In all three modes of presentation, each subject was asked to read through the material presented, and to refer to it as much as he pleased in answering the following questions:

(1) What factors do you consider were of most importance in bringing about the union between St Ninian's and St David's?

(2) Do you think that this union will be permanent?

(3) Do you think that this union will be harmonious?

(4) Is it likely that the fisher-folk will remain loyal to the united church?

(5) How would you explain the rapid formation of an enthusiastic group of followers round MacKeen, and the subsequent disintegration of this group?

It was carefully pointed out that if any subject found himself unwilling to express a definite opinion he should say so.

III. ANALYSIS OF RESULTS OF CASE I

(1) *Numerical analysis*

I propose first to give, in Tables I–III, a simple numerical analysis of the answers to questions (2), (3) and (4), and then to attempt to classify, and to determine the relative emphasis on, the leading conditions on which the various types of answer to all five questions have been based.

Table I. *Will the union be permanent?*

Group		Yes	No	Conditional	Total
English:	Academic	17	2	6	25
	Non-academic	9	2	4	15
Scottish:	Academic	11	9	—	20
	Non-academic: Clerical	15	2	3	20
	Lay	19	—	1	20
	Total	71	15	14	100

Table II. *Will the union be harmonious?*

Group		Yes	No	Conditional	Unwilling to express opinion
English:	Academic	10	5	10	—
	Non-academic	3	3	—	9
Scottish:	Academic	4	12	1	3
	Non-academic: Clerical	12	3	5	—
	Lay	16	—	4	—
	Total	45	23	20	12

Table III. *Will the fisher-folk remain loyal to the united church?*

Group		Yes	No	Conditional	Unwilling to express opinion
English:	Academic	—	12	6	7
	Non-academic	—	6	9	—
Scottish:	Academic	—	17	3	—
	Non-academic: Clerical	1	9	7	3
	Lay	—	9	11	—
	Total	1	53	36	10

Note. Of the 'Noes', the following in each group predicted loyalty as long as the new minister remained, then final lapse: 8, 4, 5, 0, 5.

With regard to these three tables, the following points seem worth noting:

(1) No subject in any group found himself unwilling to hazard an opinion as to the permanence of the union, although 14 % altogether give conditional answers—25 % of the two English groups and 6·6 % of the three Scottish groups.

CENTRAL LIBRARY.

On harmony, the total percentage of conditional answers goes up to 20 %, and 12 % find themselves unable to express an opinion. This 12 % is made up from two groups only, in the following proportions—60 % of the English non-academic and 15 % of the extremely 'canny' Scottish academics.

In Table III, the percentage of conditional answers goes up to 36 %. The percentage of those unwilling to express an opinion remains about the same, but is made up from two different groups—English academic and Scottish clerical.

These figures seem to bear out very forcibly Bartlett's contention that, "when an ordinary person is confronted with a concrete description of a social situation...he can usually give a considered opinion as to its probable upshot" (*loc. cit.*). Indeed, in view of the paucity of data on which to frame an answer to questions (3) and (4), it is surprising that more subjects did not refuse to answer. Conditional answers are, of course, almost as valuable as a direct affirmative or negative, for they generally indicate what leading conditions the subject has in mind, although particular determining factors are considered absent.

(2) While 71 % think permanence likely, only 45 % predict harmony. In other words, 26 out of 71, or 37 % of those who predict permanence, evidently consider that a certain amount of disharmony will not destroy the union. Of this 37 %, the 50 % who, by giving conditional answers, entertain the *possibility* of disharmony, thereby admit that disharmony and a permanent union are compatible.

(3) Each of the five groups is so small that further use of the material may significantly alter the results so far obtained. With this reservation, the following comparisons are made:

Table I. The distribution of opinion is most similar in the English academic and Scottish clerical groups. The greatest disparity occurs, not between any one of the English groups and any one of the Scottish groups, or between the two English groups taken together, and the three Scottish groups taken together, but between two of the Scottish groups—Scottish academic and Scottish lay. In the former, 55 % predict permanence, and 45 % breakdown; in the latter, 95 % predict permanence, and none breakdown. Whether this disparity has any significance or not will be discussed when the grounds on which these different opinions are based come to be analysed.

Table II. On the question of harmony, the Scottish groups are much more decisive than the English, 40 % of the English academic giving conditional answers, and 60 % of the English non-academic refusing to

answer at all. Taking the two English groups together, 32·5 % predict harmony, and 20 % disharmony. In the three Scottish groups, 53·3 % predict harmony, but on the other hand 25 % think disharmony likely. This is entirely due to the optimism of the non-academic groups being offset by the peculiar pessimism of the academics.

The answers to this question, of course, unlike those to the question of permanency, are conditioned by the connotation which each subject gives to the word 'harmonious'. None the less, it is interesting to note that the subjects who have had most experience of situations of this type in real life are most decisive in their answers. One would have expected that acquaintance with the number and complexity of factors which determine social 'harmony' would have engendered greater caution.

Table III. The solitary cleric who thinks it probable that the fisher-folk will remain loyal to the church does so because "St Ninian's is the church of their fathers; the endowed church, the church of the poor in past times, making no great demand on their pockets and attachment to it being comparatively easy". None of the rest of the clerical group considers that the fisher-folk will remain loyal even to the new minister, although this opinion is expressed by 25 % of each of the other two Scottish groups, and by about 30 % of each of the English groups. The Scottish academics (85 %) are most emphatic in predicting a final lapse, while the opinion of the other four groups is rather similar—from 40 to 48 % in each considering permanent loyalty unlikely.

(2) *Qualitative analysis*

The leading conditions, on which the various answers to all five questions were based, are classified in Tables IV–XII. This classification unfortunately robs the answers of all their 'individuality', but on the other hand seems to give a fairly good indication of the relative emphasis placed on each type of factor both by all five groups taken together, and by each group taken separately. The method of scoring employed was simply to count the occurrence of any factor whatsoever in any answer as one (1 unit), to classify it, if possible, under an appropriate type, and then to add up the total number of 'units' of each type. This explains why the number indicating the emphasis placed by a group on reasons of a certain type is sometimes greater than the total number of individuals in that group. In his "Preliminary Report and Suggestions" (*loc. cit.*) Bartlett determines the emphasis on any one type of factor simply in relation to the total number of individuals in the group in question. When the method of scoring just described is employed, however, the most accurate index seems to be obtained by calculating the emphasis in relation to the *total number of reasons mentioned by the group.*

(i) *Initial union.*

Table IV gives the classification of, and relative emphasis on, the factors selected as being of importance in *bringing about* union in the first instance. Table V shows the relative emphasis, calculated as described above (i.e. in relation to total number of reasons mentioned by each group separately, and by all five groups together).

In several of our experiments on 'social thinking' we had noted the recurrence of a 'conventionalized principle' which might be expressed as follows—"It is not so difficult to preserve an agreement between two

Table IV

	Group					
	English		Scottish			
Factor	Academic (25)	Non-academic (15)	Academic (20)	Clerical (20)	Lay (20)	Total (100)
I. By uniting, both congregations at least suffer no loss of prestige, and in some respects flatter their self-esteem	17	20	10	34	30	111
II. Union benefits both churches:						
(*a*) Economically	9	7	4	10	5	35
(*b*) 'Spiritually'	—	—	—	6	—	6
III. Natural 'conservative' tendency to unite against disruptive factors	8	2	11	—	3	24
IV. Enlivening and cathartic effect of recent turmoil	3	—	25	1	9	38
V. Decline of interest in religion generally (union a desperate measure to survive)	4	—	—	—	1	5
Total no. of 'factors'	41	29	50	51	48	219

Table V

	Group					
Factor	E.A. %	E.N.A %	S.A. %	S.C. %	S.L. %	All 5 groups Av. %
I	42	69	20	67	63	51
II (*a*)	22	24	8	20	10	16
(*b*)	—	—	—	12	—	3
III	20	7	22	—	6	11
IV	7	—	50	2	19	17
V	10	—	—	—	21	2

groups as to bring them together in the first instance; and a breakdown of the preliminary negotiations is more likely than a breakdown of the final agreement". An opportunity to comment on the peculiar chain of events leading up to church union in Thorness was therefore given, in the hope that further light would be thrown on this interesting generalization.

The outstanding characteristic of the replies thus obtained is that the majority of them indicate a further and much more fundamental principle. Explicit reference to this principle *as a principle* is made by only two subjects. It is most clearly enunciated by one of the Scottish clerical group: "Groups are seldom, if ever, capable of disinterested conduct, and the co-operation of groups depends on how far the interest of the group is furthered thereby and the pride of members in their own group, which flatters their self-esteem, satisfied." In nearly every answer, however (except those given by the Scottish academic group), the specific influences selected as being of most importance in the events leading up to union are such that their full significance is revealed only when the principle itself is discerned, and they are seen in relation to it. It seems reasonable to infer, therefore, that subjects who make this implicit or unconscious use of the principle have had, in real life, empirical evidence of its working, and are able to recognize factors relevant to it, although they may never have consciously formulated the principle as such. (In Aristotelian language, they have perceived the 'that', but not the 'why'.) If this conclusion is justified, our enquiry into the way that people think about social situations is not only revealing the nature, frequency and distribution of opinions, but is also bringing us very close indeed to the real causes which govern social harmony and discord.

The term 'interest' in the principle as enunciated is not intended to refer only to *economic* interest, but to include all 'gains'. In order to bring out clearly that the 'prestige-gain' was considered much more important than the purely economic benefits of union in this instance, I have given each a separate heading in Tables IV and V; and this is further justified by the fact that only the economic gain is explicitly mentioned in the text. The 'benefit of being spiritually invigorated', if one might call it such, occupies a rather peculiar position and will be discussed later.

The specific circumstances which I consider to have been picked out, consciously or unconsciously, as concrete manifestations of the 'prestige-principle' at work, are perhaps worth mentioning in detail:

(a) St Ninian's have 'asserted their independence'; by uniting, and 'saving the church'; can return without *loss of dignity*; are returning to *their own* church; will be friendly to St David's, and not stress their claim to priority, because St David's *came to them*; stragglers will return once they see the church is a 'going concern', association with which gives social standing.

(b) St David's have declined gradually, not catastrophically; are 'sacrificing their independence to save St Ninian's'; have a major share in the appointment of the new minister, after the *voluntary* retirement of their own.

(c) The new minister is 'neutral'.

No one subject, of course, mentioned all these factors. Detailed analysis reveals, however, that they are stressed most evenly in the Scottish clerical and lay groups, and in the English non-academic group. The two academic groups pick out as most important the fact that the new minister is 'neutral'—a fact which the clerics themselves completely ignore.

Some notes on the distribution of emphasis can now be conveniently made (Table V):

(a) The emphasis on the 'prestige-principle' is strikingly similar in the three non-academic groups. It is considered rather less important by the English academic, but is still the factor which they emphasize most. Only in the Scottish academic group does it occupy a subordinate position. The probable reason for this will be seen when we discuss Factor IV). Taking all five groups together, we see at once that the 'prestige-principle' is by far the most important.

(b) Considerable weight is placed on the economic gain by the two English groups and by the clerics; the Scottish academic and lay agree in roughly halving this emphasis.

(c) That the clerics alone should stress the benefits of spiritual invigoration is interesting. It should be pointed out, however, that their terminology—"the people's joy in seeing their church full, and in being carried forward on the tide of a deep enthusiasm", etc.—admits of a cynical interpretation in terms of the prestige-principle. Perhaps that would be unfair. But there is one fact which is not brought out in Tables IV and V, and which is important in this connexion. 20 % of the clerical group, and about the same percentage of the total number of subjects, go out of their way expressly to state that the causes which facilitated union in the first instance, and on the basis of which permanent co-operation can be predicted, are very far removed in tone and motive from those outlined in the Moderator's speech. Perhaps this interpretation is not meant to be extended beyond the facts of this particular case, and may be due to the manner of its presentation. None the less, the point is made very forcibly in the form of a generalization by one of the clerical group—"If you surcharge a congregation with an emotion like love or hate, or any variety of either, you will get it to unite with anyone or anything from 'Mammon' to the 'father of lies'. I have not heard of two congregations, in the Presbyterian period, who united on the basis of loyalty to the Lord Jesus, brotherly love, or social happiness."

(*d*) A natural conservative tendency to unite against disruptive factors is most clearly discussed by the two academic groups. This tendency is traced to the preponderating influence of the 'solid elements in the community', who will desire to resume the even tenor of their way after the recent emotional outburst, and will be glad to wrap themselves up once again in the 'comfortable conformity of organized religion'.

This opinion may be due either to the description of the "temper of the people" in the text, or to the fact that the majority of those who express it are academic psychologists, or to both these causes.

(*e*) What I have called the 'enlivening and cathartic effect' of the recent turmoil is stressed to a remarkable degree by the Scottish academic group. The dispute over MacKeen is considered "to have shaken the people out of their lethargy, and to have let them work off 'emotional steam'"; to have surcharged them with emotion, and so fitted them for union (cf. (*c*) above); to have quietened the die-hards, who in ordinary circumstances would have held things up; to have humbled many by showing them how they have been 'duped'; to have engendered all-important *individual* friendships, and to have accelerated the swing back towards orthodoxy once the inevitable reaction had set in.

It will be noted, in Table VIII below, that many of those who considered the union to have been thus 'emotionally engendered' also consider, for the same reason, that it will not be permanent.

(*f*) It would be interesting to discover the prevalence in other groups of the generalization implicit in Factor V, which is considered to be of some importance only by the English academics. The generalization might be expressed thus: "Any two groups who fulfil much the same function in a community and who, while vigorous, would have resisted all attempts at union, will co-operate immediately that they are so reduced in numbers or vitality as to be faced with destruction."

(ii) *Permanent union.*

Table VI gives the classification of, and relative emphasis on the grounds on which a *permanent* union is predicted; and Table VII shows the same emphasis in percentages.

Table VI

	Group					
Grounds	E.A. 17/25	E.N.A. 9/15	S.A. 11/20	S.C. 15/20	S.L. 19/20	Total 71/100
I. Relationship to superior groups	7	—	2	—	4	13
II. Factors relieving early 'stresses and strains'	20	15	12	18	19	84
III. Economic benefit	5	7	4	14	6	36
IV. No conflict *in principle*	1	—	4	10	18	33
V. Stabilizing factors which will soon begin to work	11	9	9	15	11	55
VI. Disinterested desire for welfare of the Church	—	—	—	7	—	7
Total no. of reasons	44	31	31	64	58	228

Table VII

	Group					
Grounds	E.A. %	E.N.A. %	S.A. %	S.C. %	S.L. %	All 5 groups Av. %
I	16	—	7	—	7	6
II	46	48	38	28	33	37
III	11	23	13	22	10	16
IV	2	—	13	16	31	15
V	25	29	29	24	19	24
VI	—	—	—	11	—	3

As might be expected, there is a considerable overlap in Tables IV and VI, for many of the causes which facilitated union in the first instances are also thought to be causes which will make for permanence. The economic benefit is obviously considered to have binding as well as attractive power; and in those factors which are classed together as "relieving early 'stresses and strains'", and which receive greatest emphasis in every group, it is possible to discover the 'prestige-principle' once again, in slightly different form. The absence of factors damaging to the self-esteem of either group, and the presence of some likely to flatter it, is noted; but most subjects derive their confidence, in predicting permanency, not from knowledge of this fact merely, but from the fact that the new minister is *tactful*, knowing the people's 'touchy points', and therefore likely to keep in abeyance factors damaging to their self-esteem. The Scottish clerical and lay groups make most of this point, but it also receives considerable attention from the English academics.

What was called, in Table IV, the 'cathartic effect' of the recent dispute also comes under heading II, in Table VI, for a few subjects in every group credit the people of Thorness with some genuine repentance (at least on the part of the prodigals) and a desire to make amends.

This relief of early stresses and strains is evidently considered important as it gives the various stabilizing factors (classed under V) a chance to begin their work. These factors are ranked second in importance by all groups except one, and will be discussed in more detail later.

'Relationship to superior groups' takes two forms:

(*a*) The position of the united church as a sub-group within a larger organization with central control is stressed by the English academic and by the Scottish lay groups, but in different ways. The English academic group are evidently utilizing a more or less 'theoretical' generalization; the Scottish group point, somewhat bitterly, to the fact that the Church authorities have the power, through the imposition of 'economic sanctions', to force congregations to unite, and to make any attempt at disruption almost impossible; and quote examples to show that, in spite of the promise in the Moderator's speech, this power has already been exercised.

(*b*) The English academic group, in common with the Scottish academic, consider also that the authority of the Presbytery, the body in immediate control, has been enhanced by its tactful handling of the MacKeen case. Hence the emphasis on factors of type I by the English academic group.

(3) It is worth noting, as an interesting sidelight on the method of scoring adopted, that the emphasis on type II factors is relatively low in the two Scottish groups, not because these factors are mentioned *less often* by these groups, but because other factors are mentioned *more often*.

(4) Economic benefit is ranked third in importance by all five groups together, and (strangely enough) receives most emphasis, relative to other factors, in the clerical group.

(5) The Scottish groups point out that "interdenominational antipathies, as between the former U.F. and Established Churches, though still real, are faint, and not based on any *principle*". It is therefore confidently expected, especially by the lay group, that they will, in this instance, fairly soon disappear.

(6) The 'stabilizing factors' (V) correspond in some degree to the 'conservative tendency' noted in Table IV. Some subjects say baldly that 'time and habit' will bind the two congregations together; others are more specific. "People of the Thorness type being naturally conservative", will, it is thought, stick to the habit of attending the united church, once that habit is acquired, and will prefer the "solidarity of organized religion" to any further attempt at disruption. Pride in the new group will be quickly engendered, and the two groups will boast of "getting on together better than they ever thought possible", and the fine church-building of St Ninian's will become a "centre of approving sentiments". The rise of a new generation who have never known the old antipathies will finally seal the union.

(7) Only the clerics do the people of Thorness the credit of suggesting that they *may* be actuated by a disinterested desire for the welfare of the Church. Over against this we must set the statements noted above, in Table IV, II (*b*).

(iii) *Impermanent union.*

Table VIII

	Group					
Ground	E.A. (2)	E.N.A. (2)	S.A. (9)	S.C. (2)	S.L. (0)	Total (15)
I. Union in reality externally imposed, and as such cannot resolve fundamental differences in religious temperament and outlook	2	—	4	2	—	8 (27%)
II. Embittering effect of coercive action by Church authorities	—	—	—	2	—	2 (7%)
III. Chastening effect of MacKeen incident will wear off	—	2	14	—	—	16 (53%)
IV. Disturbing effect of presence of third church (St John's)	—	—	—	2	—	2 (7%)
V. Little real religious feeling behind union	—	—	2	—	—	2 (7%)
Total no. of reasons	2	2	20	6	—	30

Table VIII shows the grounds on which breakdown is predicted. Only two reasons are given any weight, and the others are stated in detail merely to show the type of factor which tends to be introduced the more closely subjects are acquainted with the situation, or kind of situation, presented.

The remarkable emphasis placed by the Scottish academic group on III above is the natural corollary of their explanation of the event of union in the first instance. The causes which brought about co-operation being, in the opinion of this group, essentially emotional and transitory, they predict that the chastening effect of the recent turmoil will wear off before the two congregations have had time to settle down; that old and new jealousies will then appear, fostered by 'sticklers' and 'die-hards'. In this instance it would be necessary to know only the number and influence of these in order to determine just how soon the inevitable split would occur.

The other ground on which breakdown is predicted is a generalization which one of Bartlett's 'descriptions' ("Strange Effects of Drought", *op. cit.*) had already elicited—viz. that religious differences, being deep-seated and the outcome of inescapable temperamental characteristics, cannot be resolved by any externally imposed union. Being "inherent in human nature", these temperamental variations have in the past "disrupted great bodies of men into parties; and when, after a period of years, the parties try to come together, or are united by others, the antipathies or differences in temperament make themselves evident and resist easy union". I quote the more emphatic of the two clerics who predicted breakdown, and who both, as one might expect, base their generalizations on a study of Scottish ecclesiastical history.

The Scottish academic group, who consider the effect of the recent dispute too superficial to uproot old jealousies, or to prevent the appearance of new ones, do not trace these to any essentially psychological cause, but merely state that they will be resuscitated and fostered by the more conservative and cantankerous elements in the community. It would be interesting to know how far this group, and popular opinion generally, would be prepared to accept the generalization quoted above as an explanation of the undoubted fact that such jealousies and antipathies occur with great frequency in church politics.

The conditional answers do not require to be tabulated. In order to give a definite answer about 60 % of the fourteen subjects would require to know whether the new minister "is really fit to handle the united church"; the other 40 % would require to know the number and probable attitude of influential people in each congregation taken separately. On the whole, however, all these subjects seem to tend towards 'permanency'.

(iv) *Harmonious union.*

Table IX shows an analysis of the causes considered likely to make for *harmonious* relationships; it is practically self-explanatory, in view of the comments which have already been made on Table VI. The general opinion, most marked in the Scottish clerical and lay groups, is

Table IX

	Group					
Cause	E.A. (10)	E.N.A. (3)	S.A. (4)	S.C. (12)	S.L. (16)	All 5 groups (45)
I. Early stresses and strains relieved by:						
(a) Character and tact of new minister	5	3	1	10	10	29 (39%)
(b) Emotional purge of MacKeen incident and some genuine repentance	4	2	2	2	4	14 (19%)
II. Once difficult first period past, risk of disharmony diminishes owing to:						
(a) Stabilizing factors	3	—	2	4	6	15 (20%)
(b) Weakening of interdenominational antipathies (no conflict in principle)	1	—	1	6	5	13 (18%)
(c) Absence of class and economic distinctions	2	1	—	—	—	3 (4%)
Total no. of reasons	15	6	6	22	25	74

that the character and tact of the new minister will be the main determining factor in relieving early stresses and strains. Once the difficult first period is past, risk of disharmony will decrease owing to the various stabilizing factors which also make for permanency, and which have already been discussed.

The only marked difference between the English and Scottish groups is that the former place some faith in the absence of marked class and economic distinctions, while the latter utilize again their special knowledge of church history to emphasize the absence of any conflict *in principle.*

(v) *Union with disharmony.*

Table X gives an analysis of the causes considered likely to produce disharmony. This table also requires little explanation, as there is a

Table X

	Group					
Cause	E.A. (5)	E.N.A. (3)	S.A. (12)	S.C. (3)	S.L. —	Total (23)
I. Differences in outlook and temperament	2	2	1	2	—	7 (20%)
II. Chastening effect of MacKeen incident wears off, and risk of disharmony increases	2	2	18	1	—	23 (65%)
III. Sea-board communities proverbially distrustful and quarrelsome	—	—	2	—	—	2 (6%)
IV. No real religious feeling behind union	2	—	1	—	—	3 (9%)
Total no. of reasons	6	4	22	3	—	35

marked similarity between it and Table VIII. The outstanding point of difference is the increased emphasis on the probable recurrence of conflict once the superficial causes which are supposed to have removed it begin to lose their effect, and every slight disagreement is exaggerated "as such disagreements always are in small and otherwise dull communities".

Of those who refuse to be definite, the Scottish groups specify only one leading condition on which they consider harmony to depend—the attitude and number of influential people, and of 'die-hards', in each congregation; the English academic group consider social harmony is too dependent on particular circumstances even to admit of a prediction; while the English non-academic group give no reason for not expressing an opinion.

(vi) *The fisher-folk.*

The question of the loyalty of the fisher-folk was introduced in order to see how definite an answer can be elicited from a minimum of data. The relevant data in the text consist of three statements—that the fisher-folk are poor, and form almost a separate community within the town; that they are not "great church-goers, but are prone to short-lived bursts of intense religious enthusiasm"; that they were attracted first by MacKeen, and again by the first minister of the united church, who has himself an interest in the sea. As will be seen, the phrase "short-lived bursts of intense religious enthusiasm" seems to have been most heavily 'weighted' by every group of subjects; and to have encouraged considerable elaboration in setting and argument.

Reference to Table III (p. 216) shows that of the fifty-three subjects who think a relapse on the part of the fisher-folk into their usual 'Sabbath lethargy' almost inevitable, twenty-two think that the fisher-folk are likely to remain loyal to the church as long as the new minister is in charge. Every one of the twenty-two gives as his reason the new minister's 'interest in the sea', and evident sympathy with and understanding of the fisher-folk's way of life. The Scottish lay group suggest rather more subtly that their common interest will form an *exclusive* bond between minister and fisher-folk, and so encourage the latter to overcome their self-consciousness in a predominantly middle-class church. None of the clerics considered loyalty to their hypothetical colleague at all likely.

Reasons for predicting a gradual lapse on the part of the fisher-folk into non-church-going habits are given in Table XI. This table brings out clearly the emphasis on the phrase "short-lived bursts of religious

Table XI

	Group					
Reasons	E.A. (12)	E.N.A. (6)	S.A. (17)	S.C. (9)	S.L. (9)	Total (53)
I. Fisher-folk's instability in religion	10	4	12	9	7	42 (59%)
II. Individualistic and conservative outlook	3	2	7	4	2	18 (25%)
III. Poverty (compared with 'bourgeoisie' of Thorness)	1	2	—	4	4	11 (16%)
Total no. of reasons	14	8	19	17	13	71

enthusiasm". Most of the answers go further, and endeavour to explain this instability, and the conservatism of the fisher-folk, as being for the most part due to their way of life, which "produces a temperament and a cast of mind wholly unsuited to consort with a typically urban population". The difference is more than one of economic organization, but the first step towards securing co-operation between two such groups would have to be a change in economic organization—in this instance, the bringing of the fishermen into some kind of 'Fishermen's Union', so as to encourage the abandonment of their highly individualistic and conservative outlook.

Every one of the thirty-six subjects who gives a conditional answer makes the issue turn on the ability and methods of the new minister—will he be able to control class-jealousies and save the pride of both fisher-folk and 'bourgeoisie', and will he be willing to adopt 'MacKeen methods' in preaching? One or two incalculable factors are also suggested—the presence or absence of revivalists, the poverty or prosperity of the fishermen, the possibility of a change in their economic organization, the degree of inter-marrying between fisher-folk and townspeople, and so on. Three of the Scottish lay group suggest ingeniously that, in view of the professional's scorn of an amateur, the minister's interest in the sea may prove a doubtful advantage!

(vii) *The influence of a leader.*

Table XII gives the reasons brought forward to explain the rapid formation of a group round MacKeen.

In the English academic groups 'MacKeen qualities' receive 50% of the emphasis; in the non-academic groups approximately the same stress is laid on the fact of opposition. Otherwise the relative emphasis in each group is so similar as to indicate either that reasons of the type used are strongly suggested by the narrative presented, or that there is

Table XII

Reasons	Group					Total
	E.A. (25)	E.N.A. (15)	S.A. (20)	S.C. (20)	S.L. (20)	(100)
I. Mere novelty	6	4	6	3	4	23 (9%)
II. MacKeen's qualities	19	13	21	26	18	97 (39%)
III. Action of Presbytery in opposing MacKeen	11	20	19	37	26	113 (46%)
IV. Cumulative excitement	2	2	6	1	2	13 (5%)
Total no. of reasons	38	39	52	67	50	246

widespread agreement about the factors which are important in situations of this kind. Every subject certainly seemed more eager to tackle this question than any of the preceding ones.

MacKeen's qualities are generally thought to have been enhanced by contrast with the old minister, and can be listed as 'youth, personality, leadership, emotional appeal, oratory and unconventionality'. A few of the Scottish subjects think that he must have had some 'genuine message' and several of the clerics point out, perhaps a little enviously, that MacKeen's lack of academic qualifications was more than compensated for in the eyes of the people by the fact that he 'gave results'.

The effect of the Presbytery's action in opposing MacKeen is worked out very elaborately. All groups agree that the opposition, being apparently unjustified, inflamed the pride of MacKeen's followers, and encouraged them to assert their independence. MacKeen himself, on being opposed, immediately became a *rebel* and a *martyr*. Now 'a rebel *within a tradition* is always supported'—a very interesting generalization common to all groups, and evidently intended to apply to rebels 'within' and against *all* kinds of tradition—and as a martyr, 'a good man not receiving justice', MacKeen would attract many to his 'cause'. As one cleric remarks, quoting John Galt—"God's corbies are ill craws to shute at", and it was just when MacKeen, a minister of the Gospel, came to be regarded as being unjustly persecuted, that his supporters became thoroughly consolidated and ready to make sacrifices for him. An essentially psychological but somewhat cynical reason for this 'love of a cause' is suggested by several subjects both English and Scottish: such causes make an eloquent appeal to a man's egotism, and flatter his belief in his own importance.

Table XIII gives the reasons for the disintegration of the group round MacKeen.

Once again the emphasis throughout all groups is so similar as not to merit further analysis.

Table XIII

	Group					
Reasons	E.A. (25)	E.N.A. (15)	S.A. (20)	S.C. (20)	S.L. (20)	Total (100)
I. Mere novelty soon wears off	4	3	3	—	1	11
II. MacKeen's shortcomings	13	12	14	18	15	72
III. Removal of opposition	11	8	12	10	9	50
IV. Financial responsibility	5	6	6	13	9	39
V. Conservatism	10	4	3	7	5	29
VI. Loss of prestige and social status	4	—	2	—	3	9
VII. MacKeen and his church cannot give full 'pastoral services'	—	—	—	5	1	6
Total no. of reasons	47	33	40	53	43	216

The first three reasons are just the antithesis of those postulated to explain the formation of the group. MacKeen is now seen as a personality whose appeal is strictly limited, and unable 'long to oppose the weight of tradition'; while the scandal concerning him in such a small and respectable community can have but one result. The removal of opposition is taken as demonstrating the reverse of the generalization quoted above—"a rebel, though effective 'within' a tradition, becomes ineffective once he sets himself up as wholly independent". The 'cause' loses its glamour and its publicity, and an ignored martyr is merely ridiculous.

(*To be continued*)

CHARACTERISTIC MOTIVATION IN THE ACTIVITIES OF SCHOOL-GIRLS

By M. D. VERNON
From the Cambridge Psychological Laboratory

Part II

III. *Distribution of the drives in the three groups* (pp. 232–237).
IV. *Interests and occupational wishes* (pp. 238–245).
V. *Summary* (p. 245).

III. Distribution of the drives in the three groups

We have described how the various drives and motivational tendencies functioned in a number of individuals, and have indicated how it was possible to detect the existence of these drives in the responses to questioning. It is now desirable to give some idea of the frequency of these drives and the extent to which they affected the different girls in the three groups. We have already pointed out that it was easier to find examples of certain of the drives in certain of the groups than in others, which indicates that the frequency varies between the groups. If now we study the number of individuals in whom the drives appeared fairly clearly, we find a confirmation of this variation. Table I gives the number of times each drive occurred, and the percentage of individuals in each group showing these drives (the totals are greater than 100, because most individuals possessed more than one drive). These figures must not be taken too literally; it is not possible always to decide exactly what drives each individual possesses. But they give a rough idea of the frequency of distribution. We shall discuss in the next section the exact forms of behaviour in which the drives were manifested.

We observe that the frequency of certain drives was approximately the same in all three groups, while for other drives it varied between the groups. Among the former was *social conformity* (the decrease between Group (*c*) and Groups (*a*) and (*b*) was too slight to give any definite evidence of a decline in frequency). Social conformity was much the most frequent of all the drives, functioning in about three-quarters of all cases. Thus we may infer that the great majority of school-girls from 12 to

Table I

Name of drive	No. of occurrences of the drive in Group			No. of occurrences as percentage of no. in the Group		
	(*c*)	(*a*)	(*b*)	(*c*)	(*a*)	(*b*)
Social conformity	36	22	21	71	76	75
Humanitarianism	10	3	0	20	10	0
Independence	13	6	3	25	21	11
Security	6	5	6	12	17	21
Activity	19	13	10	37	45	36
Superiority	27	11	4	53	38	14
Dominance	4	2	0	8	7	0
Display	8	4	3	16	14	11
Pleasure	10	6	6	20	21	21
Total	133	72	53	262	249	189

18 will seek to make close contact with their social group, to conform to its traditions, imitate its behaviour, or obey its wishes. Any or all of these tendencies indicate the presence of the social conformity drive. In general the social group to which behaviour and ideas were referred was that of school companions. But we have also described cases, chiefly among the younger or emotionally retarded, of conformity to parents; of conformity to school staff; and, among the more mature, of conformity to general social opinion—appearing as a moral obligation towards society. Thus as age increases and emotional maturity is achieved, the circle of society to which the individual conforms becomes wider. But we have described cases in which, for various reasons, the individual is unable to conform freely to the social group to which she should naturally be suited. Emotional backwardness is one such cause; it appeared in those cases who were unable to make contacts with their contemporaries because they were too closely tied to their parents, sometimes because their parents exerted too strong a forcing or restraining influence upon them, keeping them in childish subjection. Intellectual and cultural inferiority also prevented normal contact with and conformity to school companions. Cases such as that of M 5 in Group (*c*) (this *Journal*, XXIX, 127) will inevitably occur in any social group. They were rather unduly frequent in Group (*a*) for the following reason: This group consisted partly of girls coming from highly intelligent and cultured homes in the professional classes, and partly of girls from less cultured (and sometimes impoverished) homes in the commercial and business classes. The former were naturally more intelligent and usually showed more energy, initiative and independence than the latter, and consequently tended to dominate the school group and determine its

general social and cultural level. Thus a feeling of inferiority was apt to arise in the girls of lower cultural and social status, and they were unable to conform satisfactorily to the ideas and interests of their more intelligent companions. This inability to conform resulted in shyness and timidity, inability to make social contacts, and, in some cases, the tendency to take refuge in the society of animals. Thus we have included among the cases of social conformity those in whom the drive was thwarted or inhibited, as well as those in whom it functioned naturally and freely.

Activity was also comparatively evenly distributed in the three groups, occurring in between one third and one half of the cases; it was somewhat more frequent in Group (*a*) than in Groups (*b*) and (*c*). The occurrence of this drive was difficult to decide with any certainty, except in some of the outstanding cases described in the last section. In point of fact almost all of these girls were naturally fairly active, except some who were emotionally inhibited. But we attributed the drive only to individuals who seemed capable of fairly purposive and prolonged effort; and not to those who were indeed 'active' in the usual sense of the word, but flitted rapidly from one pursuit to another, with little concentration or perseverance. The more extreme cases of this lack of directedness and perseverance were usually motivated by the drive towards *pleasure* or *variety*. We observe that the relative frequency of this drive was almost identical in the three groups. Thus this type of laziness or superficiality seemed to be an inherent characteristic which varied comparatively little with age or social background.

Coming now to the drives which varied in frequency, we see that humanitarianism, independence, superiority, dominance and display were most frequent in Group (*c*) and least frequent in Group (*b*); while security was most frequent in Group (*b*) and least frequent in Group (*c*). In all cases Group (*a*) was intermediate—an interesting point, since the girls in Group (*a*) were intermediate in age and social status between those in Group (*c*) and those in Group (*b*). However, we cannot necessarily infer from these figures that the drives towards humanitarianism, independence, superiority and dominance were almost, or altogether, absent in Group (*b*). It is possible (1) that these tendencies were too slight to appear in the behaviour of these girls, which was largely governed by conformity to the social norm of the group; and (2) that even when the drives did affect behaviour, most of the girls were too shy and inarticulate to express them in any way in their very brief descriptions of interests, wishes, etc. But against this latter supposition it must be remembered that the girls' behaviour in school was also observed, and was reported on by their teachers. Thus we may perhaps conclude that certain of the drives do appear more frequently and clearly among girls who are superior in age, intelligence, culture and social status. Which of these factors is the most important cannot always be certainly determined; that almost certainly varies from one drive to another.

To take first the drive of *humanitarianism*: It was pointed out in the previous paper (*Brit. J. educ. Psychol.* (1937), VII, 302 and (1938), VIII,1) that an interest in people for themselves was comparatively rare among younger girls, and did not appear as a rule until they were nearly adult. The results given in this paper seem to bear out this conclusion. We see that there were few cases except in Group (*c*) of girls with a genuine interest in or concern for people in themselves—as apart from those who simply liked being with people or working with them. The latter we have distinguished as being motivated mainly by social conformity. But a further difficulty arises here. There were in addition, in all three groups, quite a number of sociable and friendly individuals, with not much independence or initiative, who seemed to show sympathy and protectiveness—manifested, for instance, in a wish to take up work with little children; B 2 in Group (*c*) (see this *Journal*, XXIX, 126) was a case in point. But it was very hard to distinguish among these cases the genuine humanitarian from the rather childish and unintelligent social conformer, since the latter might wish to do work with children, partly because it seemed to promise her less intellectual strain, and partly because in undertaking such work she could hope for social approval. Thus we have only assigned the humanitarian drive with any certainty to those girls with sufficient independence and initiative to have adopted this interest in people spontaneously, not merely to have followed social tradition. This is not to conclude that there is a necessary positive correlation between independence and humanitarianism (so that because there was more independence in Group (*c*) there was therefore more humanitarianism); but merely that we can detect the latter with more certainty when the former is also present. At a later age it might be possible to decide which among the social conformers were also genuine humanitarians; many of those who wished to take up nursing or child nursing, or even work with animals, might turn out to be such.

There is little doubt that *independence* was more frequent in Group (*c*) than in Group (*a*), and more frequent in Group (*a*) than in Group (*b*). Age may again have been partly responsible for this; girls may develop more independence when they have passed adolescence. Favourable and unfavourable selection were probably more important. In the Group (*c*) school, the girls with less independence and initiative might have left immediately after taking the School Certificate, and not stayed on in the sixth form. Among the less well-to-do school population, the girls with more drive and initiative, as well as intelligence, obtain special places at a secondary school such as that from which Group (*a*) was drawn. Those with the least intelligence and initiative remain in a Senior School, such as that which contained Group (*b*). But more important in Group (*a*) than the influence of selection was that of home background. Almost all the girls with most independence were the daughters of academic or professional class parents. Thus probably their independence was to some extent inherited; and also every sign of it was approved of and encouraged by parents.[1] It is probable that the same influence was at work in some of the girls in Group (*c*), although its effects were here less clearly marked. In Group (*b*), however, independence may quite possibly have been discouraged, both at home and at school.

The influences of selection and of home background were also at work on the distribution of the *superiority* drive. In the Group (*c*) school the most ambitious girls would tend to stay on at school in order to work for the University and for a profession. Ambition and desire to rise in social status would also lead girls to go to a secondary school, such as that of Group (*a*), and not remain at a Senior School as in Group (*b*). Indeed, we saw that two girls in Group (*b*) who did show the superiority drive were trying to go to the Technical School (they were too old to go to a secondary school). It was also true that the daughters of professional class parents such as those in Group (*a*) and Group (*c*) were on the whole the most ambitious, especially in desiring academic success and professional careers. Here indeed parental encouragement and expectations were sometimes so strong as to have a definitely unfavourable influence, since the daughter endeavoured to maintain a standard of performance which was too high for her to achieve without undue strain. A further, less deleterious, influence seems to have been at work with the Group (*c*) girls. In addition to those coming from professional homes were a large number from well-to-do and upper class homes; these girls were accustomed to a superior social position, and to succeed in obtaining almost anything for which they wished. Thus it was natural that they should expect this success and their superiority to the rest of mankind. This probably accounts for the large number, more than half the group, in whom the superiority drive was observed.

---

[1] An exception to this was H 2 in Group (*a*) (see this *Journal*, xxix, 133); but she had inherited so much independence from her father that it merely became more aggressive when it came in conflict with his commands.

Although the relative amount of *dominance* was greater in Groups (*c*) and (*a*) than in Group (*b*), the actual number of cases was too small to suggest any definite conclusions. It appears that this drive is infrequent even in intelligent girls of superior social status. Probably it is an inherent quality of the individual, not very common in women, and relatively unaffected by social status. But it is also possible that it develops and appears more clearly with increasing age; it occurred in more than a third of the University students (*loc. cit.*), who came from much the same social classes as those of Groups (*c*) and (*a*).

The frequency of *display* did not vary greatly in the different groups, though the largest proportion of cases appeared in Group (*c*). It seems probable that in every social class there are a certain number of individuals who are chiefly interested in occupations and pursuits in following which they may hope to be personally admired. At the same time, this quality may well have been encouraged by the more fashionable and well-to-do parents of the girls in Groups (*c*) and (*a*).

Finally, we see that *security* alone of all the drives was relatively most frequent in Group (*b*) and least frequent in Group (*c*). It may be supposed that this is in the main a directly environmental effect. Poverty, unemployment and inadequate nutrition were only too familiar among the Group (*b*) girls. In Group (*c*) such factors were almost entirely absent except in the two or three cases described in Part I of this paper whose fathers were in relatively insecure financial positions. At the same time it is interesting to observe that the security drive appeared in Group (*b*) more in a general timidity of manner, a desire to cling to parents and to familiar home surroundings, than in the wish for "a good, safe job" (as in Groups (*a*) and (*c*)). The security given by earning a good wage in a job with little fear of unemployment was outside the experience and beyond the expectations of the Group (*b*) girls; and thus security was constituted by the support of home and parents. On the other hand the Groups (*c*) and (*a*) girls were beginning to realize that one can obtain security for oneself, and is not obliged to depend for it on other people.

This concludes the general account of the distribution of the drives; and we shall next consider how they manifested themselves; and observe that the actual interests and occupations to which they lead also varied greatly from one group to another.

IV. INTERESTS AND OCCUPATIONAL WISHES

Since those who were strongly motivated by social conformity endeavoured to imitate the behaviour and adopt the wishes and interests of their social group, it was natural that behaviour, wishes and interests should vary according to the social norm of the group—the traditional and customary ideas, interests, wishes and behaviour—which itself varied considerably with the social setting. In Group (*b*), the group of lowest social status, the norm was the narrowest in range, giving the fewest permitted variations; it consisted chiefly of domestic and practical activities. The main interests were in cooking and needlework in school, and in simple amusements such as reading and knitting out of school. The desired occupation was domestic service. The horizon was bounded by the town nearby and its openings for employment and amusement. Opportunities for physical activity were distinctly inadequate, though singing, physical training, and playing games at school (on a small scale) were usually popular, and sometimes cycling, walking and swimming out of school. Manual activities also provided an outlet for the activity drive. Interest in people was confined mainly to the familiar friends and acquaintances in the school and village; though the wish to do child nursing may have indicated a nascent humanitarianism. The social conformers were fairly contented with this life; but those who possessed a strong security drive tended to shrink away from it, and to wish never to go away from home but live always with their parents in the same familiar places. On the other hand, those who possessed some degree of the superiority drive wished to rise above their social level to a higher one. This could be done by (1) going into a shop or a hairdresser's establishment, or (2) going to the Technical School, training as a shorthand typist, and going into an office—or better still, becoming a school secretary. The first course was suitable for those girls who possessed comparatively little academic intelligence, but some social flair. It also appealed to those in whom the drive towards display existed. The second course was chosen by those with greater academic intelligence. It was probably regarded as the superior of the two since it was likely to lead to a better paid job.

We have already pointed out that Group (*a*) formed in fact a divided social group, one section of girls coming from the professional classes, the other from the business and commercial classes. The latter in the main conformed fairly closely to their social norm, which was not very different from that of the Group (*b*) girls although they were some degrees above them in prosperity and in social status (the intermediate social classes would go to technical and central schools). This social norm was constituted by wishes and interests which were mainly practical and domestic. It is true that the wider school curriculum gave opportunities which were sometimes adopted of making wider cultural interests—as, for instance, in history, English literature, science, drawing and craftwork; although spontaneous interest in these subjects seemed somewhat inhibited by the feeling of inferiority arising from their more moderate culture and academic ability. Out-of-school interests in reading, knitting, needlework, walking, cycling and swimming were the same as those of Group (*b*), but playing games, riding, gardening, keeping pets, drawing and music were added, and gave a fuller and more varied scope for activity than was possible for Group (*b*). Occupational wishes included: secretarial work (4), domestic science (1), nursing (3), work with children (3), work with animals (5).[1] Domestic science and work with children are the counterpart, at a higher social level, of the domestic service (cook and nursemaid) of Group (*b*). That is to say, the Group (*a*) girls also desired in the main work of a practical rather than an intellectual nature, which involved plenty of contact with people. The wish to do nursing and work with children possibly indicated the existence of the humanitarian drive; but as was pointed out above, this was not certain. The wish to work with animals was probably a phantasy wish, created by the desire for social security arising in a situation of social insecurity and inferiority. For other girls, security was represented by "a good job", "something settled", something in which one can soon earn a living wage. But this desire to be a wage earner was also in some cases a mark of the superiority drive, as, for instance, in the case of H 5 (this *Journal*, XXIX, 139). Cases of superiority were, however, comparatively rare in this section of the Group (*a*) girls. There were a certain number who possessed the drive towards display, which manifested itself principally in a fondness for dancing and acting.

The social norm for the professional class section of the girls in Group (*a*) was wider, and the permitted variations greater. It was predominantly intellectual in character. School interests were imaginative and cultural, often literary; and three girls had strong scientific interests. Out-of-school interests were numerous and varied, and included artistic and literary pursuits as well as physical and manual activities. Moreover, these interests were usually pursued with more activity and greater energy than by the girls previously described. Occupational wishes included: University careers (7), teaching (4), medicine (1), writing (1), art (1), music (1), the better type of secretarial work (3).[1] It seems probable that the social tradition (supported by parents and teachers) was towards some kind of profession, though secretarial work would be included if as a private secretary or organizer. (The dominance drive also directed people towards this kind of work, as shown by B 2 and T 1.) Nursing seems sometimes to have been suggested but was firmly rejected by at least two girls. The professional careers were in the main those involving work with people. But so many of these girls possessed strong superiority drives; thus they desired in addition (1) a successful University career, (2) an occupation which is "interesting", "not monotonous", something in which "you can use your brains"—in fact, something which gave opportunities of exercising their imaginative and intellectual abilities. But opportunities for social contacts were also considered important, and humanitarianism seemed to appear in one or two of those who, for instance, enjoyed history because it "was about people", and wanted to take up medicine, teaching, etc. Security did not seem to trouble this set of girls.

---

[1] Some of these were alternatives.

Although a certain number of the girls in Group (*c*) came from the professional classes, the social norm was determined in the main by the large number of girls coming from well-to-do homes, whose parents were often fashionable 'society' and 'big business' people. Thus their out-of-school interests were often of the opulent, sporting, games-playing, dancing and theatre-going variety. But they also possessed a wide range of cultural interests. The school curriculum was a varied one, and not all of these girls were doing specialized work; many were carrying out a generalized programme. There were opportunities for studying the usual school subjects, and, in addition, philosophy, economics and politics; for artistic subjects—drawing, modelling, painting, architecture, music, elocution and acting; and for domestic science. Thus cultural and artistic interests were not confined to the more intelligent and original members of the group, but also characterized the social norm to which the rank and file conformed. Particularly noticeable also, throughout the group, was an interest in foreign languages and in travelling abroad. Thus 41 % of all the girls were interested in modern languages; 51 % had been abroad in their holidays; 37 % intended going to live abroad for a period of some months, when they left school; and 16 % wanted to take up some occupation in which they could use their languages. This interest seemed characteristic of the social norm in Group (*c*); it resulted from the greater wealth and wider social opportunities and experiences of the girls in this group.

---

[1] Some of these were alternatives.

When we examine the other occupational wishes, we find it hard to determine the nature of the social norm in that respect. It is true that almost all the girls said they wanted a career. But among the less intelligent and independent this wish seemed to be largely a function of the activity drive; that is to say, a definite occupation would give the girl something to do, while if she lived at home without one she might get bored with little or nothing to occupy her. Some thought of taking up nursing or nursery school work (7), though it was not at all certain if they would persevere in these occupations. Others (8) wished for secretarial work, or to go into business. The widespread interests in the arts and crafts were mainly cultural, and did not correspond to natural abilities and inclinations of such magnitude as to make their owners wish to develop them professionally. Table II shows the number of girls interested in these subjects and the number who hoped to take them up as a career, which is very small by comparison, except in the case of architecture. The choice of architecture may have resulted from a special interest in the subject peculiar to the school, and inculcated by good teaching. But it is probable also that it was adopted because it seemed more likely to provide a secure means of livelihood than would any other artistic profession. The security drive also came clearly into operation in some cases whose parents were not well-to-do, and directed them towards relatively secure occupations such as teaching or the Civil Service, and away from the insecurity of, for instance, acting and writing.

Table II

Subject	No. interested in subject	No. taking up subject as a career
Drawing, modelling, craftwork, etc.	19	3
Architecture	4	3
Music	21	1
Elocution and acting	16	1
Domestic science	17	2

We have already shown that over half the girls in Group (*c*) were motivated by the superiority drive, and one quarter by the independence drive. The majority of these girls hoped to have a University career and subsequently to enter a profession. Thus in general they were specializing in one or more subjects in school, and they possessed strong academic interests (usually in English, history, modern languages or science), in addition to the more generalized cultural interests we have described above. The main occupational wishes were: teaching (4), medicine (5), secretarial and business (4), Civil Service (Treasury and Departmental classes) (5), research work (5), social work (5), writing and journalism (5).[1] The Civil Service, writing, research work, were chosen particularly by those who wished to demonstrate their superiority in excelling in their University and subsequent careers, and who also desired to exercise some independent initiative. The same motives appeared in many of those who were taking up medicine and were determined not to become general practitioners when qualified, but to obtain some superior position, out of the general ruck (in surgery, tropical medicine, welfare or clinic work). Journalism seemed a good opening to those with a rather superior and detached general interest in human activities, affording plenty of opportunities for independent initiative, and also plenty of change and variety. Social work and teaching appealed to those with a feeling of moral obligation towards society. This was in part the outcome of a genuine humanitarian feeling for people. But in some cases it also seemed to derive from social conformity—the desire to obtain social approval for a type of work which also gave opportunities for dominance and superiority (see, for instance, M 4, described in this *Journal*, XXIX, 136). Secretarial work seemed to be somewhat of a *pis aller* to the individual who possessed strong superiority and independence drives. She regarded it as an occupation more suitable for her less gifted companions, but one which it might be necessary to adopt if other more interesting careers were impossible; thus A 2 said that "if she fails (to get into the Civil Service), will be the inevitable secretary".

In cases in which there was a drive towards display, there was often great interest in some form of artistic self-expression, such as acting, elocution or dancing, with a wish to take it up professionally. When there was also a strong superiority drive, the individual desired to excel in these professions, or else rejected them altogether, for fear of not excelling. Another occupational wish probably arising from these drives was to go into Parliament. But sometimes display was linked with the drive towards pleasure and variety. In such cases the individual usually wanted to act or dance as a pastime or a part-time occupation, and not as a full-time career. The pleasure seeker was in general characterized less by any particular set of interests and occupational wishes than by a general changeableness, a desire for variety and a lack of perseverance. This might result, as in the case of S 1 in Group (*b*) in the mainly passive behaviour of staying at home and reading novels; or it might appear as the desire to "have a good time", with plenty of games, amusements and travelling; with plenty to do, but no obligation to work hard either in a career or in housework carried on at home.

---

[1] Some of these were alternatives.

In Table III are summarized and outlined the interests and occupational wishes in which the various drives appeared to manifest themselves in the three groups. Group (*a*) is subdivided into (1) those coming from non-professional class homes, (2) those coming from professional class homes. The activity drive has been omitted, since this seemed in general to reinforce the carrying out of activities the nature of which was determined by other drives.

We see that, just as some drives varied greatly in frequency of distribution in the different groups and others very little; so the manifestations of some drives varied more between groups than did those of others. Clearly the manner in which social conformity appeared varied as the social norm varied. The superiority drive showed a like variation, at least in occupational wishes, since its possessor wished to rise superior to the social norm. Display directed the individual towards some occupation which was admired in that particular social setting, varying from that of hairdresser to that of Member of Parliament. Security led to a retreat from the unfamiliar to the most familiar (the home) in Group (*b*); but in Groups (*a*) and (*c*) it also led to positions of economic security or social escape (work with animals). Humanitarianism varied less, though even there in the higher social classes the interests were more academic and the occupations mainly professional. But independence naturally led its possessor to seek for something independent of social tradition, in which originality and initiative could function; and the Arts seemed to form the most obvious vehicle for these. Dominance required roughly the same type of social setting throughout (a leader and subordinates) for its manifestation, though its actual context varied; and indeed opportunities for its exercise were very slight for those of lower social status. Finally, pleasure varied a little in context but was always characterized by changeability and a general lack of effort and perseverance.

Again, it is clear that any one interest or activity may be motivated by very different drives, even in the same social group, and still more in different social groups. Thus teaching may be desired through (1) social conformity, (2) humanitarianism, (3) security, (4) superiority, or (5) dominance, or any combination of these drives. Artistic activities which are normal to the social conformer in one group may be original (independent) or superior for individuals in other groups, or provide opportunities for pleasure or display for others. Clearly then we are not justi-

Table III

Drive from which interests and wishes derive	Interests				Occupational wishes			
		Group (*a*)				Group (*a*)		
	Group (*b*)	(1)	(2)	Group (*c*)	Group (*b*)	(1)	(2)	Group (*c*)
Social conformity	Domestic, manual, physical* activities	Domestic, manual, physical* activities	Intellectual, artistic, physical,* manual activities	Domestic, artistic, manual, physical* activities; humanistic interests; foreign languages and travel	Domestic	Secretarial, nursing, work with children	Professional, higher secretarial	"A career"; secretarial, nursing, work with children
Humanitarianism	?Playing with children	?History ?Playing with children	?History ?Science	Zoology; history, economics, politics, philosophy	?Nursemaid	?Nursing, work with children	Teaching, medicine	Social work, teaching, medicine
Independence	—	—	Original work, especially artistic and literary	Original work, especially artistic and literary	—	—	Artistic and literary work; "something interesting"	Artistic and literary work; journalism; research
Security	Staying at home	Doing things with animals	—	Dislike of going away to school	Staying at home	Secretarial; work with animals	—	Living at home work with animals; teaching, Civil Service
Superiority	Arithmetic, English	—	Intellectual or academic subjects, especially those at which she excels	Specializing in academic subjects	Technical school, shorthand typist; hairdresser, shop assistant	Being a wage earner	University career, professional success	University career, Civil Service, writing, research
Dominance	—	—	School duties	Prefect and games duties	—	—	Secretarial, organizing	Social work, teaching
Display	Needlework	Dancing, acting, games	Acting	Dancing, acting, physical activities	Shop-work, hairdressing	Dancing	—	Dancing, acting beauty culture M.P.
Pleasure	Novel reading	Games, amusements, travel; anything new	—	Games, amusements, travel	Not going out to work	Freedom from parents and domestic work	Something interesting, not monotonous	Having plenty to do, but not a regular job

* 'Physical activities' include: games, sports, physical training.

fied in predicting anything as to the nature of interests, or of the behaviour or occupational wishes to which they lead, unless we know something of (1) the basic drives of the individual who possesses these interests, and (2) the social setting and background from which he comes and in which he is situated.

V. SUMMARY

1. From the answers to questions given by three groups of girls of different ages and social settings, was demonstrated the existence of the following drives towards: social conformity, humanitarianism, activity, independence, security, superiority, dominance, display, pleasure or variety.

2. The frequency of distribution (that is to say, the proportion of individuals in whom the drive appeared clearly) was roughly the same in the three groups for social conformity, activity and pleasure. In the group containing the girls of highest age and social status the drives towards humanitarianism, independence, superiority, dominance and display were the most frequent, and security the least frequent; and the position was reversed in the group containing the girls of lowest age and social status. The intermediate group was intermediate in frequency for these drives. The various factors which produced this effect are discussed.

3. The particular interests and occupational wishes towards which the drives directed individuals differed in the different groups, particularly in the case of social conformity, superiority, display, security and humanitarianism. Moreover, any particular interest or occupational wish might be produced in different individuals, in different social groups, by very different drives. Hence the occurrence of any such interest or wish has no significance unless considered in conjunction with the basic drives of the individual (and also probably his age and emotional maturity), and with his social background and environment.

In conclusion, I wish to make acknowledgement to the Medical Research Council for a personal grant which made it possible for me to carry out this work. My thanks are due to Prof. Bartlett and to others for whose advice and suggestions I am most grateful; and most particularly to the Headmaster, Headmistresses and Staff of the three schools in which these observations were made, for their kindness and helpfulness.

(*Manuscript received* 11 *May* 1938)

A PRELIMINARY NOTE ON A VOCATIONAL INVESTIGATION

By ALICE HEIM

I. *Introductory* (p. 246).
II. *Sources and type of data* (pp. 246–247).
III. *Results* (pp. 247–251).
  (1) *Employability* (pp. 247–249).
  (2) *Choice of work* (p. 249).
  (3) *Personal qualities* (p. 250).
  (4) *Criteria of satisfaction* (pp. 250–251).
IV. *Summary* (p. 251).
V. *Conclusion* (p. 251).

I. INTRODUCTORY

In this paper, a preliminary survey is made of some of the points which have arisen in the course of an enquiry into the working of the present system by which elementary school-boys enter industry. The Industrial Health Research Board has kindly allowed the investigator a grant, enabling her to carry out the work.

II. SOURCES AND TYPE OF DATA

The main aim of the investigation is to determine the importance of, and the relations between, a number of the factors which influence the early industrial lives of elementary school-boys. The data are drawn from four schools, three Senior and one Central. Material has been derived from the boys' school-masters, relatives, doctors, employers and employment officers, and information obtained as to the boys' scholastic ability, appearance and manner, temperament and character, physique, home environment, choice of occupation, satisfaction and satisfactoriness in work, spare time activities and employability.

The material here discussed is derived from the first of the three groups of school-leavers who are the subjects of the investigation. This group consists of ninety-eight boys, most of whom left school in March 1936. (Some of the Central School boys left later in the same year.) The investigator gave a group test to all these during their last term; this was a comprehensive test which takes half an hour to give and consists of three parts: Directions, Analogies and Opposites, and Mechanical Aptitude.

The scholastic ability, appearance and manner, and temperament and character of the boys were judged by their school-teachers and recorded on forms constructed on a three-point scale. The first form asked for information such as the boy's speed and originality of thought, imagination and consistency in school work; good and bad school subjects were also mentioned. The second form dealt with such characteristics as normal facial expression, voice, clothes, and the third listed various social, emotional and 'moral' qualities and their opposites.

Full reports on the health and home environment of each boy were obtained from the School Medical Officer and the School Enquiry Officer, respectively. The investigator also obtained some first-hand information on home conditions by personal visits to the boys' parents, when hobbies, ambitions and views on present and past situations were discussed. Data as to choice of work and work obtained were derived also from the Juvenile Employment Office.

An 'employability index' was made out for each boy on the basis of the quality of the situation (or situations) he was in during his first year after leaving school. A fivefold division was adopted, ranging from α, i.e. highly skilled occupations with excellent prospects, to γ, i.e. quite unskilled work with no prospects. When a boy had had more than one situation, an attempt was made to estimate his 'average' employability; in these cases, the upward or downward trend of a boy's career was also taken into account.

III. RESULTS

(1) *Employability*

Fourfold tables were constructed to show the relations between various pairs of the above factors, and the probability of obtaining the observed relations by chance was calculated in each case by the χ^2 method.

The following tables show the connexion between employability and test scores (I), scholastic ability (II), home environment (III), and appearance and manner (IV).

It will be seen that the relationship is in each case statistically significant. It is interesting that the connexion between employability and the test scores should prove to be stronger than that between

Table I. *Relation between employability and test scores*

	Test interquartile groups	
Employability	1 and 2	3 and 4
α, $\beta+$ and β	26	13
$\beta-$ and γ	13	24

$\chi^2=7{\cdot}5$. $P=0{\cdot}006$.

Table II. *Relation between employability and scholastic ability*

	Scholastic interquartile groups	
Employability	1 and 2	3 and 4
α, $\beta+$ and β	34	22
$\beta-$ and γ	13	25

$\chi^2=6{\cdot}36$. $P=0{\cdot}01$.

Table III. *Relation between employability and home environment*

	Home environment	
Employability	α and $\beta+$	β, $\beta-$ and γ
α and $\beta+$	37	3
β, $\beta-$ and γ	21	35

$\chi^2=29{\cdot}35$. $P=0{\cdot}0000001$.

Table IV. *Relation between employability and appearance and manner*

	Appearance and manner	
Employability	V.G. and G.	Av., F.O. and P.
α and $\beta+$	27	11
β, $\beta-$ and γ	20	36

$\chi^2=11{\cdot}31$. $P=0{\cdot}0008$.

employability and scholastic ability: the latter distribution would occur as a matter of chance once in one hundred, the former, only six times in one thousand.

However, according to the values of χ^2 found and given above, both home environment and appearance and manner were much more closely related to employability than were the scores for intelligence or for scholastic ability. This would suggest that the importance of the role played by certain social factors in the boys' early industrial lives, although suspected, has perhaps not been fully recognized in the past. On the other hand, it should not, in the opinion of the investigator, necessarily be inferred that a boy's home environment and appearance and manner play in fact a greater part in his industrial career than do his more intellectual abilities and attainments: even if this be true of his first year after leaving school, it seems likely that this importance will decrease as he grows older. Home influences will probably weaken during the next few years and emphasis may gradually shift from *appearance* to *experience* when vacancies for older boys are to be filled.

(2) *Choice of work*

A comparison of the boys' choices of work with those chosen for them by their parents, suggested that parents are on the whole more ambitious for their sons than the boys are for themselves. Table V shows the distribution of the vocational wishes, within the categories of 'skilled', 'semi-skilled' and 'unskilled' occupations. Office work, shop work, the Services and the railway were listed separately, as they cover a wide range of occupations with considerable variation in the degree of skill required. The figures in brackets at the top of each column give the total number of choices which were ascertained and from which the percentages in the table are taken. Thus, 32·2% of the eighty-seven boys whose first choice is on record expressed a desire for skilled work.

Table V. *Boys' and parents' choice of work*

Type of work	1st choices (87) %	2nd choices (70) %	Parents' choices (41) %
Skilled	32·2	17·1	31·7
Semi-skilled	23·0	24·3	14·6
Unskilled	9·2	18·6	0·0

It will be seen that while the percentage of parents whose choice is for skilled work is almost identical with that of the boys' first choice, there is considerable difference in regard to 'semi-skilled' and 'unskilled' occupations; not one parent who expressed any views wished his son to enter an unskilled occupation, although 9·2% of the 'first choices' and 18·6% of the 'second choices' are for such work. The hypothesis that those parents who possessed definite and explicit opinions as to their children's careers were also the most ambitious merits consideration. No evidence has been obtained on this point, but it is one which should not be ignored in any discussion of the above results.

A large proportion of parents wished their sons to go on to the railway or to do office work, the former on account of its reputation for security, and the latter largely on account of the social status which is conferred by cleanliness and shorter hours. These occupations were slightly less popular among the boys.

(3) *Personal qualities*

The investigator has called on, and corresponded with, the boys' firms in an endeavour to obtain a report on the suitability, progress and conduct of the boys in their work and also to ascertain employers' general views on juvenile labour. Employers were asked what they considered to be the most important qualities required by employees who were engaged on work such as their firm did. Responses ranged from silence to long harangues, and the subject-matter of the answers given presented an equally wide range. From the answers it appears that employers either do not know or cannot express what it is they look for in their employees. If this is the case, the high juvenile labour turnover found by the investigator in numerous firms (at a time when the number of available positions exceeded the number of available boys) is hardly surprising.

However, an examination of the relationship between temperament and character traits and employability suggests that some kind of sorting-out process, on the basis of personal qualities, nevertheless takes place. When these traits were compared with employability, it was found, for instance, that no boys who were designated as outstandingly *conscientious* have had γ situations. On the other hand, among the α situations, not a single strikingly *lethargic* boy was to be found.

Owing to the paucity of the number so far examined, these findings may be far from reliable. They suggest, however, that further research along the same lines might yield interesting results.

(4) *Criteria of satisfaction*

An analysis of the reasons given for dissatisfaction with work, and for leaving situations, indicated that the physical conditions under which boys work and the type of work which they are required to do play a greater part in determining the degree of a boy's satisfaction, and the length of time he will remain in a situation, than material considerations such as hours and wages. This conclusion is opposed to the views on the subject expressed by most employers and many Juvenile Employment Officers, who usually account for juvenile discontent and unrest in terms of insufficient remuneration.

The above finding, however, is confirmed by two facts which became evident in the course of the investigation: (i) the continuance of a high juvenile labour turnover in various firms which have recently raised their wages while continuing to ignore other relevant human factors; (ii) the existence, in several of the larger firms, of a higher labour turnover in their highly paid, mass production departments than in their less well-paid workshops, where boys are taught a skilled trade.

Among elementary school children, as in other groups, there are naturally individuals for whom high wages for short hours are the greatest attraction a situation can offer and, in many cases, economic pressure at home prohibits any consideration other than substantial and speedy remuneration. However, the assumption that this is the guiding principle of the vast majority of school-leavers and their advisers when they first enter industry would appear, from the present data, to be unwarranted.

IV. SUMMARY

The main points discussed in this paper may be summarized as follows:

(i) The relationship between employability and the intelligence test scores proved to be stronger than that between employability and scholastic ability.

(ii) The relationships between employability and both appearance and manner, and home environment, were found to be stronger than either of the connexions referred to in (i).

(iii) Parents appeared to be on the whole more vocationally ambitious for their sons than the boys themselves.

(iv) Very little agreement was found among employers as to the most important qualities required by their juvenile employees.

(v) A comparison of employability with temperament and character traits suggested that in practice the distribution of boys among good and bad grade situations may to some extent be determined by the possession or lack of certain such traits.

(vi) Remuneration was found to play a smaller part in determining occupational satisfaction than did the type of work done or the physical conditions under which it is carried out.

V. CONCLUSION

Owing to the small number of boys so far considered and the short time during which they have been employed, every conclusion is of a tentative nature. Material on similar lines concerning the second and third groups of school-leavers is being collected and the investigator is keeping in touch with members of the first group and their employers. The examination of these further data will furnish some guide as to the value of these first suggestions.

(*Manuscript received* 10 *March* 1938)

THE EFFECT OF ADAPTATION UPON VISUAL ACUITY

By K. J. W. CRAIK

From the Psychological Laboratory, Cambridge

I. *Introduction* (p. 252).
II. *Apparatus and procedure* (pp. 253–257).
  (1) *General description of apparatus* (pp. 253–254).
  (2) *Details of apparatus* (pp. 254–255).
  (3) *Calibration* (p. 256).
  (4) *Experimental procedure* (pp. 256–257).
III. *Results* (p. 257).
IV. *Discussion* (pp. 258–263).
V. *Conclusions* (pp. 263–265).
VI. *Summary* (p. 266).
References (p. 266).

I. Introduction

It was found by Mellone and Rawdon-Smith (1935, unpublished), and in a more extended investigation by the present author (1938), that visual adaptation has a considerable effect upon intensity discrimination. The differential threshold was found to be lowest when tested at the same illumination to which the eye was adapted, and to be raised considerably where these 'test' and 'adapt' illuminations were widely different.

It seemed that a similar investigation of visual acuity might throw some light on the relation between this and intensity discrimination. A positive result was to be anticipated from the experiments of Lythgoe (1932) in which acuity was diminished by a surrounding field, except when that field had approximately 1/10th the brightness of the smaller field containing the test object; for spatial and temporal proximity of an adapting field appear in general to have very similar effects.

II. APPARATUS AND PROCEDURE

(1) *General description of apparatus*

The requirements of the apparatus for this investigation may be summarized as follows. To secure adaptation there should be a uniformly illuminated field, preferably subtending more than 30° at the eye, and capable of variation over a wide range of brightness. The 'test' field should be approximately the same size in order that the retina may be exposed to the test illumination over a wider area than that covered by the test object. Otherwise, glare and 'contrast' effects may disturb the acuity (Lythgoe, 1932). The test field should be variable over the same range of brightness as the adapting field, but independently of it. Within the test field should appear an object with parts whose angular separation is variable from 0 to 40 min. Its black portions must be as completely dark as possible, and the ratio of the widths of the black and white portions should remain constant while the separation varies. The 'adapt' and 'test' fields must have the same spectral composition and be rapidly interchangeable without requiring movement of the subject's head or eyes.

Considerable difficulty was encountered in meeting all these requirements simultaneously, and it was decided to employ a principle which rendered possible the use of high illuminations, rapid interchange of 'test' and 'adapt' fields, and convenient measurement of widely different acuities, though it entailed a field subtending only 16°.

The light from an electric bulb, rendered nearly parallel by a condenser, filled a large lens at whose principal focus the eye was placed. This lens appeared uniformly illuminated by the 'Maxwellian view'. Close to the condenser was a test object consisting of two parallel strips of thin, blackened metal, joined at their ends so as to form a rectangular frame with a central gap, equal in width to one of the strips. This frame could be rotated through 90° about its longer axis, so that the strips and the gap appeared to become narrower, if viewed from a fixed position. A small lens of short focus, mounted between the condenser and the large lens, cast a much reduced image of the strips (and of the bright background formed by the condenser) in front of the large lens, so that the image was seen sharply focused in the centre of the large field. The small lens was moved out of the optical path in the 'adapt' position, and a double filter-frame near the eyepiece, operated by a key-arm in conjunction with the small lens, enabled the illumination to be altered.

The appearance of the 'test' field as seen by the subject is shown in Fig. 1.

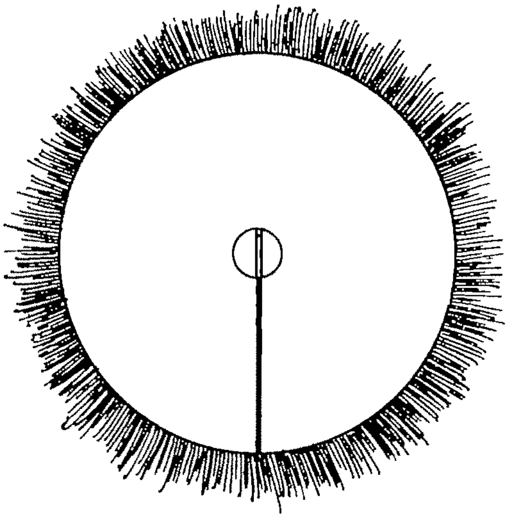


Fig. 1. Appearance of field with test object in centre.

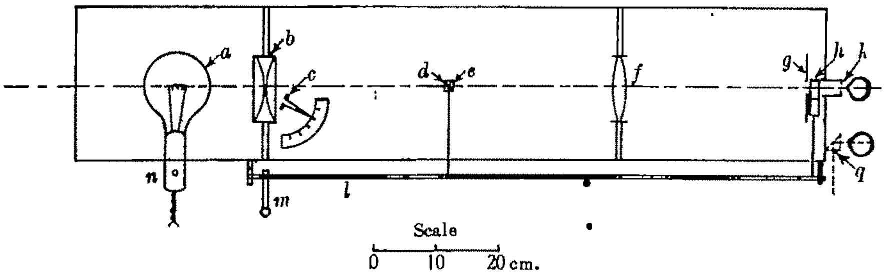


Fig. 2. Diagram of apparatus. *a*, 100 W. pearl bulb. *b*, condenser. *c*, test object and scale. *d*, prism, and *e* lens, both in movable frame. *f*, field lens. *g*, stop. *h*, double filter frame. *k*, artificial pupil. *l*, rod. *m*, key-arm. *n*, switch. *q*, prism for calibration.

(2) *Details of apparatus*

A condenser *b* (Fig. 2), 10 cm. in diameter, is placed in front of a 100 W. 200 V. 'pearl' bulb *a*. Immediately on the other side of the condenser is the test object *c*, consisting of two blackened metal strips 10 cm. × 6 mm. × 0·2 mm., fastened together, with a 6 mm. gap between them, by means of cross-strips at their ends. This rectangular frame is pivoted at one end and fitted with a pointer and scale. It can be rotated by a slow motion gear and knob; any back-lash in the latter is without effect on the calibration as the pointer is rigidly attached to the test object.

60 cm. from the condenser is a lens *f*, 10 cm. in diameter and 24 cm. in focal length,

which receives light from the condenser and, when the eye is placed at *k*, its principal focus, provides a uniformly illuminated field, 16° in diameter.

32 cm. from the condenser is a smaller lens *e* having a diameter of 10 mm. and a focal length of 10 mm., which casts a much reduced image of the test object between it and the large lens *f*. The test object *c* had to be placed slightly off the optical axis in order to prevent it from casting a shadow over the Maxwellian field of *f*. To restore the image of the test object to the centre of the system, a 6° prism, *d*, 10 mm. in diameter, was mounted in the same cell as the lens *e*. This cell, constructed of metal 0·2 mm. thick, was mounted on a metal arm presenting only a thin edge, 1 mm. wide, in the direction of the optical axis. These precautions were taken to avoid the presence of large black areas in the test field, since these would interfere considerably with acuity in neighbouring regions. The arm is pivoted at its base so as to enable the small lens and prism to be swung out of the way.

This small lens has two functions. First, it introduces an image of the test object into the field of view, without the mechanical inconvenience of moving the object and its scale. Secondly, it reduces the apparent size of the test object about tenfold, permitting the employment of a relatively large and robust object concurrently with short optical paths and compact design.

This image is unavoidably magnified about 2½ times by the lens *f* which provides the large field, but it is still considerably smaller than the apparent size of the test object at the same distance. The condenser *b* has also two functions—to concentrate light on the lens *f*, and to provide a uniformly illuminated background for the test object. The distances of the lenses are so arranged that the bright background of the test object has very nearly the same brightness as the surrounding field, and so that the apparent distance of the test object is infinity. The apparent length of the test object is 2°, and, at maximum separation, the centres of the strips subtend 14′ at the eye; this does not represent the lowest measurable values of acuity, which are nearer 30′, but it enables acuity to be measured continuously from its keenest value down to 5% of this; and the further reduction to 1% of its original value is not essential for the present experiment.

A stop is fitted at *g* to minimize scattered light, and a double filter-frame, *h*, is mounted on a pivoted lever. Each half of this frame holds two Wratten neutral filters, 25 mm. sq., mounted in B glass, or alternatively one filter and a clear glass square, or a blank stop. An artificial pupil 3 mm. in diameter is fitted at *k*, in front of the subject's eye.

A rod *l*, rotated through a small angle by a key-arm *m* moves the filter frame so that either half is in the optical path, and at the same time swings the lens and prism (*d* and *e*) in or out of the field. Thus, in the 'adapt' position, a uniform field of 16° is exposed; on depressing the key to the 'test' position, the other set of filters becomes operative and the small lens swings to the centre of the field, exposing the test object. The links actuating the filter frame can be detached when equal test and adapt illuminations are to be employed, thus avoiding duplication of filters. The change from 'adapt' to 'test' position occupies about 0·1 sec.

A switch at *n* enables the light to be extinguished while the sets of filters are changed between series of readings.

The optical system is mounted inside a light-tight box having ventilating holes and light traps. The whole is mounted on an adjustable table, and fitted with a chin-rest for the subject.

(3) *Calibration*

The calibration falls into three portions:

(1) The angles subtended at the eye by the field, the small circle containing the test object, and the angle between the lines of the test object at maximum separation, were measured in the following way. A right-angled prism was fixed to the box at *q*, and a screen with a millimetre scale placed three metres distant. This permitted binocular comparison between the field and the scale; slight difficulties owing to the different distances of the objects were counteracted by numerous measurements. Calculation based on this method gave results in good agreement with the magnitudes computed from the distances and focal lengths of the lenses.

(2) The maximum angular separation of the test strips being thus found, the scale of *c* could be calibrated if the relative separations for different positions of the pointer on *c* were known. Although this could be calculated from the sine of the angle of rotation of the pointer, this was considered inadvisable owing to the finite thickness of the strips and the possibility of slight mechanical inaccuracies. A small lens, similar to *e*, was placed between *e* and *f*, the lens *f* being temporarily removed. This additional lens cast a greatly enlarged image of the image of the test object on a screen with a millimetre scale, placed at *g*. By means of the pointer, the test object was set to give the distances, in millimetres on this scale, corresponding to the required angles of separation, and the pointer-scale marked accordingly.

(3) The calibration of the field brightness was accomplished by a selenium-oxide photo-cell of the generator type. This was clipped over the artificial pupil and the reading of the microammeter noted; this reading was then reproduced by illuminating a magnesium-oxide screen, subtending 16° at the cell, against a background of darkness. The artificial pupil was still fixed to the photo-cell. This diffusing screen was illuminated by a 1000 c.p. projector lamp, calibrated in turn by a 20 c.p. N.P.L. substandard lamp. The illumination on the screen was calculated from the inverse square law. The maximum illumination of the field was found to be approximately 10,000 equivalent ft. candles (the unit used by Lythgoe (1932) etc.). Similar measurements were made with the 1/10th and 1/100th transmission filters in position; for lower illuminations the transmissions of the filters were accepted at their assigned values. The minimum intensity was 0·0001 e.f.c., obtained by two 1/10,000th transmission filters in series.

Accurate calibration of the absolute brightness of a 'Maxwellian view' is somewhat difficult, but the above method gives sufficient precision for this investigation.

(4) *Experimental procedure*

The subject was seated with his head fixed in a chin rest, in such a position that the field appeared to him most uniformly illuminated, and so that the beam of light from the apparatus fell symmetrically upon the pupil of his right eye, which was then exposed to the 'adapt' field for a period sufficient to ensure nearly complete adaptation to the illumination employed. This varied from 5 min. for adaptation to the highest intensities, to 20 min. for dark adaptation. The test object, with the required illumination of the whole field, was then exposed for periods of 2 sec., its separation being increased or decreased progressively between each exposure; the subject indicated whether or not he was able to resolve the lines. Between these exposures adaptation was restored for 1 to 5 min.; the longer periods were necessary for the recovery of scotopic sensitivity after exposure to the highest illuminations. These periods, as well as those for preliminary adaptation, were found empirically to ensure a sufficiently steady state, without prolonging the experiment so far as to cause undue fatigue. No particular fixation point was provided; the subject was instructed to fixate the centre of the field and to do his best to resolve the test object in the time provided. Trials at equal 'test' and 'adapt' illuminations with longer exposures showed that 2 sec. was sufficient for reliable judgment. It appeared that foveal and slightly eccentric fixation were automatically employed at those illuminations at which each was most effective.

Two experienced subjects (R.S.S. and R.C.O.), both of whom had slight myopia, were employed in the main experiment, and two others (M.D.V. and K.J.W.C.), both emmetropes, in trials at equal test and adapt illuminations.

III. RESULTS

In Fig. 3, the acuity of these four subjects is plotted linearly against the logarithm of the illumination. In Fig. 4, log acuity is plotted against log illumination, for equal test and adapt intensities. Log acuity is given as log $1/\theta$, where θ is the angle in minutes subtended by the centres of the strips. This method was used by Schlaer (1937). In Fig. 5, log acuity is shown against the logarithm of the testing illumination, for different values of adapting intensity; in general it is highest where the two illuminations are equal. At 1/1000th of the adapting illumination, acuity is extremely poor—approximately 0·1—and the same at 10,000 e.f.c. unless the eye has been adapted to a higher illumination than 100 e.f.c.

The effect of large pupil size on acuity is shown in Fig. 6, in which acuity at equal 'test' and 'adapt' illuminations is shown with the normal pupil, and after dilation to 6 mm. diameter with Euphthalmine, a mydriatic which has little effect on accommodation. There is a slight decrease in acuity at the highest illuminations.

IV. Discussion

First of all, some particular features of the results may be noted. The curve relating acuity to the logarithm of intensity appears to be approximately a straight line for illuminations between 0·001 and 10 e.f.c. (see Fig. 2), thus supporting Lythgoe's results (1932). From 10 to 10,000 e.f.c., the curve is flatter and seems to be quite level above 100 e.f.c. This disagrees with the data of Lythgoe (1932) and Schlaer (1937) using a **C**-shaped test object; it accords, however, with the results

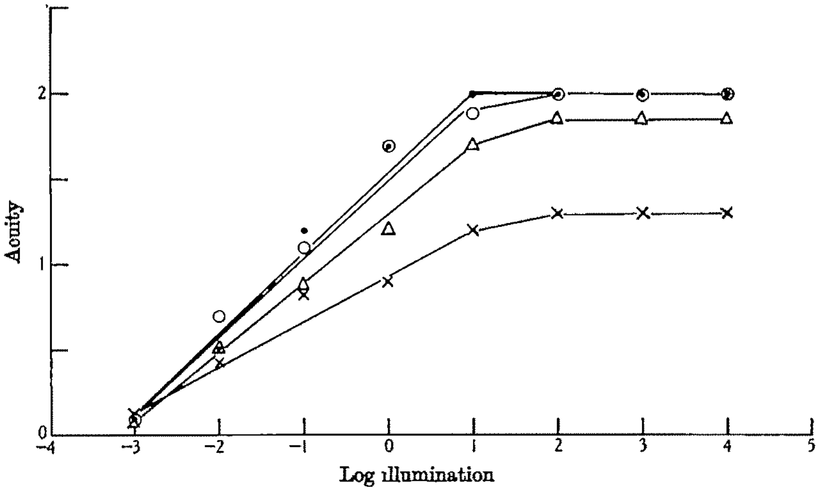


Fig. 3. Acuity against log illumination. Four subjects.

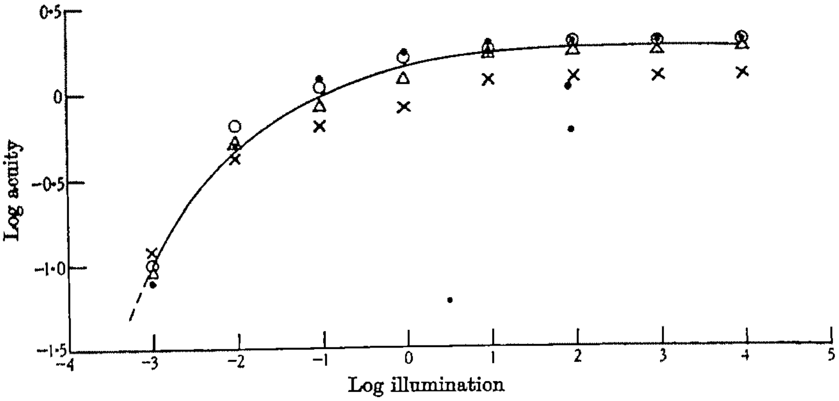


Fig. 4. Log acuity against log illumination; same data as Fig. 3.

of Banister, Hartridge and Lythgoe (1927), and of Lythgoe (1932) for a grating. Thus it seems that the discrepancy is more likely to be due to the

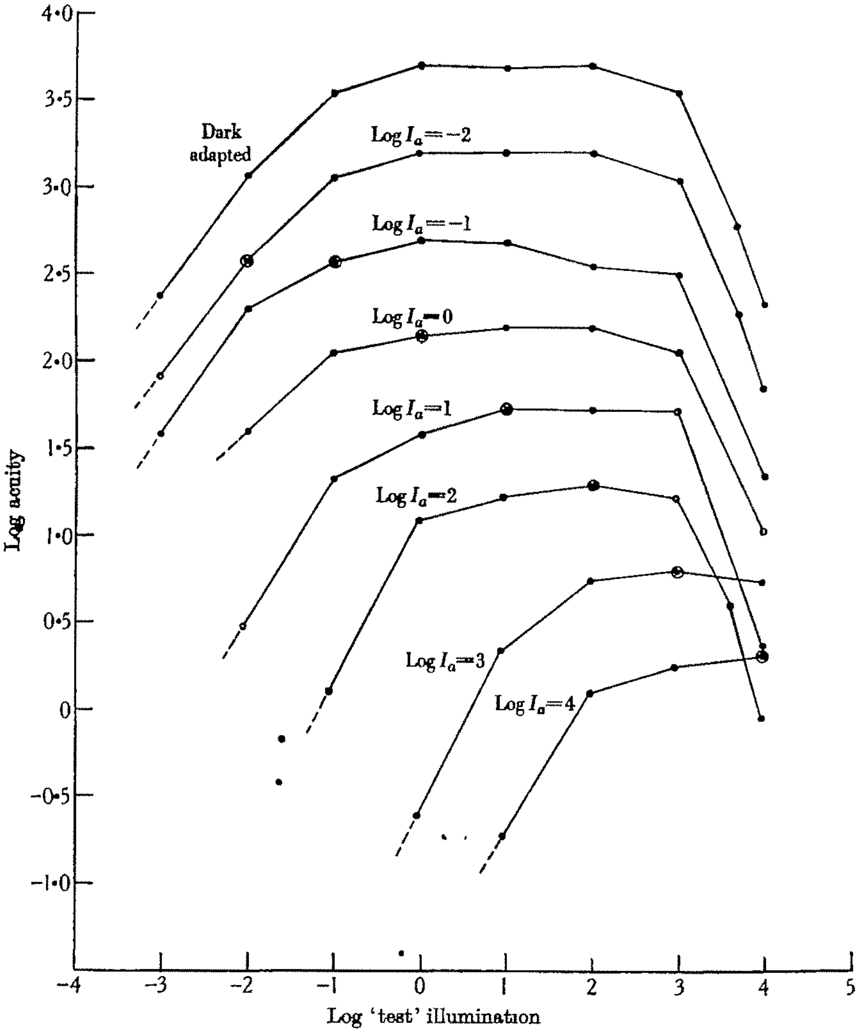


Fig. 5. Log acuity against log 'test' illumination, at various adapting intensities, marked beside the curves. Circles represent equal 'test' and 'adapt' illuminations. All curves above the lowest are shifted upwards by 0·5 on the ordinate scale, to avoid confusion.

test object employed than to any cause such as insufficient size of adapting field. Very inadequate adaptation would certainly cause such a deterioration of acuity, as the results of the present experiment confirm by means of successive adaptation; but the agreement of these data and those of Schlaer, who employed a 30° field, suggest that the 16° field used in the present experiment was adequate, and that the curve is indeed flat for a grating or parallel-line object.

Secondly, the present results neither confirm nor disprove the existence of a photopic-scotopic inflexion, since the maximum width of the test object was insufficient to give points on the curve at the requisite very low illuminations.

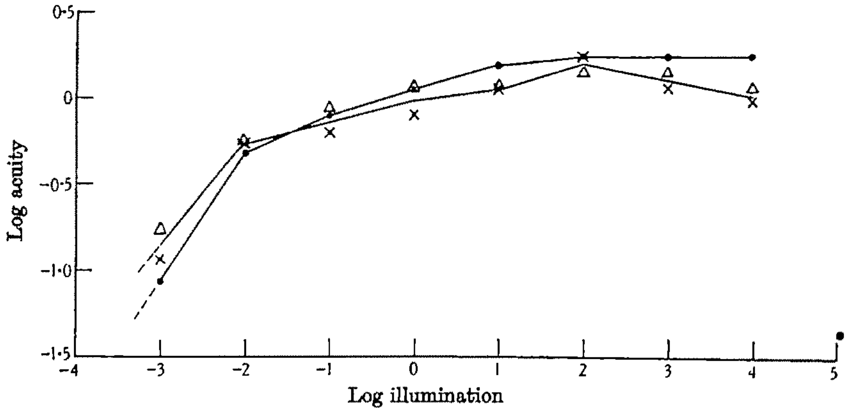


Fig. 6. Log acuity against log illumination before and after dilation of pupil with Euphthalmine. Dots, normal pupil, and artificial pupil 3 mm. diam. Crosses and triangles, two different occasions after Euphthalmine.

There are two main theories of acuity—those of Hecht (1928) and Hartridge (1918, 1923). The first endeavours to account for the effect of illumination on acuity by supposing that the sensitivity of individual rods and cones is normally distributed. Acuity will then be impaired at very low illuminations since few fresh receptors are available with the appropriate sensitivity to enable them to come into operation. This theory fits ill with the fact, demonstrated in Hartline's electrophysiological studies of nerve impulses in the vertebrate eye (1938), that the 'on fibres' give a response which increases in frequency with intensity of stimulation. Further, if the theory were correct, it might be expected that acuity would be extremely good at high illuminations after dark adaptation; for all receptors would have their maximum sensitivity, and the greatest number would therefore be available at any given illumination. The present results show, however, that under such conditions acuity is extremely bad. It is also found that if the space between the two lines of a test object which is just below the resolving power of the eye be filled in, for a certain length, this part appears much darker than the rest of the object. This can be seen if Fig. 7 be observed from a distance of about 9 ft. (The same principle is of course used to produce greys in half-tone prints.) Now if two strips are resolved if, and only if, a stimulated receptor falls between their images, this is hard to explain; for two unresolved lines must have no stimulated receptor between them, and the presence or absence of black filling should make no apparent difference. If, on the other hand, resolution is due to discrimination of intensity gradients, the phenomenon is easily explicable.

Hartridge's theory, on the other hand, offers a most attractive explanation of the way in which acuity under optimal conditions and illumination exceeds what might be expected from the aberrations of the dioptric system of the eye; but it makes no attempt to explain the effect of varying the illumination on acuity and does not appear able to do so. For according to Hartridge's calculations, the illumination gradients on the retina which correspond to the edges of objects do not exceed 5 μ or 70 sec. of arc in effective width, and no further separation of the strips beyond 5 min. of arc (from centre to centre) can increase the brightness differences on the retina. Yet it is well known that under very weak illumination separation up to 30 min. is effective in securing resolution.

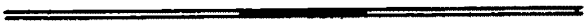


Fig. 7.

Again, relative localization of two points on the retina does not appear to diminish greatly with decreasing illumination. It is possible under such conditions to provide good localization by a framework of faint fixation points without either improving or impairing acuity.

The present experiment provides four main results:

(1) Acuity is not dependent simply upon physical illumination but also upon the state of adaptation of the eye. Thus at 10,000 e.f.c. acuity may be 2 if the eye is adapted to that illumination or less than 0·07 if it is dark adapted; while at 10 e.f.c. acuity may be 1·9 if the eye is adapted to 10 e.f.c. but 0·07 if it is adapted to 10,000 e.f.c.

(2) Acuity appears to be highest when the 'test' and 'adapt' illuminations are equal, except when the adapting illumination is less than 10 e.f.c. Lythgoe (1932) found acuity to be very slightly better when the surround was 1/10th as bright as the test field, than when it was equally bright—in this case 12·6 e.f.c. This occurred almost at the illumination where a marked effect of a somewhat similar kind was found in the

present experiment, which may be explained as follows: over the range of adapting illuminations above 10 e.f.c., the highest acuity is found when the 'test' and 'adapt' illuminations are equal, and this may be taken as the general rule. At lower adapting illuminations, however, the general decline of acuity even at equal 'test' and 'adapt' intensities enters as a complicating factor. The result will be that acuity may be higher at an illumination somewhat above the adapting one; but it is usually equal or higher still when the eye is adapted to, and tested at, that higher illumination.

(3) Acuity becomes extremely low

(*a*) when the 'test' illumination is 100 times as high as the adapting one, if this latter is higher than 100 e.f.c., or

(*b*) when the 'test' illumination is 10,000 e.f.c., if the adapting one is less than 100 e.f.c.

Acuity is also very low when the 'test' intensity is 1/1000th of the adapting one, but over a wider range of adapting intensities, i.e. from 1 to 10,000 e.f.c. Let those 'test' illuminations (above and below the adapting intensity) at which acuity becomes extremely poor, e.g. 0·07, be called the upper and lower limits of acuity for that level of adaptation. Then it appears that the lower limit is proportional to the adapting illumination over a range of 4 log units, and the upper limit over a range of 2 log units only (10,000 e.f.c. being the highest adapting illumination obtainable in the apparatus). A possible explanation seems to be as follows: The progressive shift of the upper limit is probably due to adaptation of the cones; that of the lower limit to that of the rods. Perhaps the cones do not increase in sensitivity below an adapting illumination of 100 e.f.c.; and as central vision is automatically used at high testing intensities, whether the eye be previously bright or dark adapted, there would be no increase in "dazzle" nor decrease in acuity. At low test intensities, however, the eye is involuntarily moved so as to bring the image of the test object upon the retinal area most sensitive to that illumination, which contains many rods and is capable of further dark adaptation and extension of the lower limit of acuity.

(4) The present results are very similar to those obtained by the writer (1938) for the effect of adaptation upon brightness discrimination. It has been suggested above that neither Hecht's nor Hartridge's theory affords an adequate explanation of the influence of illumination on acuity, and further that the present results, showing a decrease in the acuity of the dark adapted eye at high intensities, are contrary to what would be expected on Hecht's theory.

It has been argued by Legrand (1935) that the main factor in the deterioration of acuity at low illuminations is large pupil area, which exposes the non-spherical part of the lens. Legrand claimed to have measured acuity independently of the dioptric system of the lens by means of interference fringes formed on the retina, and to have found little decrease at low illuminations. This seemed unlikely, since in most studies of acuity, including the present experiments, an artificial pupil of small size has been used; but as a further test the pupil was dilated with Euphthalmine, which increased its diameter to 6 mm. when exposed to 10,000 e.f.c. The artificial pupil was of course removed from the apparatus. A slight decrease in acuity, from 1·81 to 1·1, was found at 10,000 e.f.c. (Fig. 6), but this was not of the same order as that occurring under low illuminations or inappropriate adaptation.

V. CONCLUSIONS

Is it possible, then, to suggest any explanation based on intensity discrimination? Kravkov (1938) suggests a connexion between the two, and Schlaer applies to his acuity data the same equations which Hecht (1934) had derived for brightness discrimination, but these throw little light upon the way in which increase in the size of the test object can make such discrimination possible. On the other hand, it has been shown that increase in the size of a test patch lowers the absolute threshold, until the test patch subtends more than 20° (Reeves, 1918). The same effect has been found for the differential threshold, until the patch subtends more than 45 min. (Lasareff, 1911). Such an increase in size could, then, compensate to some extent for the effect of decreased illumination upon intensity discrimination.

It is not to be expected that the data for acuity and brightness discrimination should be identical. The relation between the width of a test area and the reciprocal of the differential threshold is not exactly linear, as may be seen from Lasareff's data. Pupil size and the form of the lens will affect the very crowded intensity gradients which may be involved in acuity more than the larger and less critical areas used for brightness discrimination. Again, however low the illumination, it is always possible to obtain a differential brightness judgment if the variable patch or intermittent stimulus is made so bright as to be above the absolute threshold, though the background be below it; whereas Lasareff's results show that there is a limit to the effective increase in size of the test area.

In Fig. 8 the classical data of König and Brodhun (1889) on intensity discrimination, of König (1897) on acuity, and the recent results of Steinhardt (1936) on intensity discrimination and of Schlaer on acuity, are superimposed. Log $\Delta I/I$, and $\log \theta - 3{\cdot}8$ (where θ is the subtended angle in minutes, taking an acuity of 1 as 1 min.) are plotted against log I. A similar superposition is shown in Fig. 9, using the present results for acuity and those of the writer (1938) for brightness discrimination. In both cases, considerable similarity will be seen. This would be of little account if any clear-cut and insuperable difference could be found; the supposed deterioration of brightness discrimination at high illuminations, and the continued improvement of acuity, have long been held to constitute such a difference. Recent work, such as that of Steinhardt, suggests, however, that the deterioration of brightness discrimination is not so marked as had been thought, and probably does not occur with test fields smaller than 5°. Again, the decrease in acuity at 1000 and 10,000 e.f.c. resulting from dilation of the pupil by Euphthalmine (Fig. 6) suggests that, normally, constriction of the pupil may compensate for a latent decrease in acuity consequent upon impaired brightness discrimination. A final point of similarity between the two visual functions is the need for very small eye movements; steady indirect fixation of the

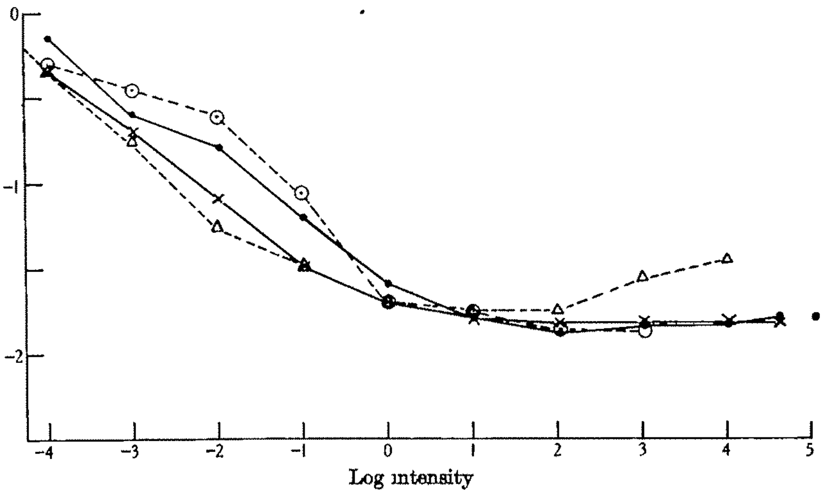


Fig. 8. Brightness discrimination data of Konig and Brodhun (broken line and triangles) and of Steinhardt (dots) and acuity data of Konig (broken line and circles) and of Schlaer (crosses), superimposed. Log $\Delta I/I$ and $\log \theta - 3{\cdot}8$ are used. To avoid confusion, only points representing integral log intensities, or the mean of the two nearest points, are plotted.

acuity test object in a somewhat similar experiment was found, under very weak illuminations, to lead to its vanishing altogether.

The effect of field size upon the absolute and differential thresholds has various possible explanations; it may, for instance, be due to some kind of retinal summation (cf. Adrian and Matthews, 1927, 1928; Elsberg and Spotnitz, 1937). A simpler mechanism might involve spatial adaptation, such as the influence of a 'surround'. A change in intensity seems to be most readily appreciated when the eye is adapted to one illumination and is in the initial stages of exposure to a slightly greater or smaller illumination (Helmholtz, 1866; Stiles, 1929; Hecht, 1935); while the work of Lythgoe and Tansley (1929) on surrounds shows a similar action to that of temporal adaptation. The effect of a surround is felt over several degrees, but diminishes with its distance from the test area. It follows that the centre of a large area of the retina, in darkness, will be more dark-adapted than the parts near a bright area, and will therefore give a greater response when, as a result of eye-movements, it is exposed to light. Conversely, those parts of the retina which are in the centre of a large bright area will give a greater decrease in response when the image of a dark strip falls on them. The difference in response between receptors, or in the same receptor at successive movements, will therefore be greater, within limits, when the areas of the bright and dark strips are increased.

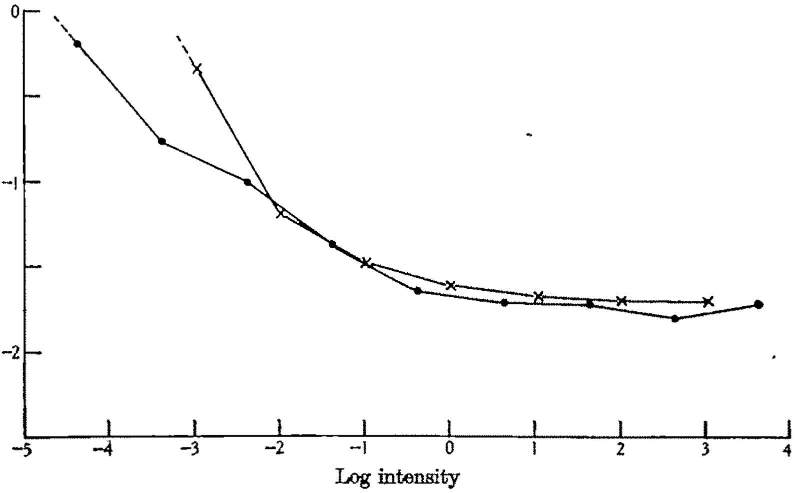


Fig. 9. The present acuity data and the writer's brightness discrimination results superimposed as in Fig. 8. Dots, Log $\Delta I/I$; crosses, log $\theta - 3{\cdot}8$.

VI. SUMMARY

Visual acuity has been tested immediately after exposure of the eye to a wide range of illuminations equal and different from those at which acuity was measured. It was found that acuity was highest (approximately 2) when 'test' and 'adapt' illuminations were equal; at illuminations far above and below the test intensity, it deteriorated greatly. The results are very similar to those previously found for brightness discrimination. A tentative explanation is proposed for these effects and for some anomalous features of them.

I am indebted to Mr R. S. Sturdy for assistance in taking readings.

REFERENCES

ADRIAN, E. D. & MATTHEWS, R. (1927). *J. Physiol.* LXIV, 279.
—— —— (1928). *J. Physiol.* LXV, 273.
BANISTER, H., HARTRIDGE, H. & LYTHGOE, R. J. (1927). *Brit. J. Ophthal.* XI, 131.
CRAIK, K. J. W. (1938). *J. Physiol.* XCII, 406.
ELSBERG, C. A. & SPOTNITZ, H. (1937). *Bull. neurol. Inst. N.Y.* VI, 253.
HARTLINE, H. K. (1938). *Amer. J. Physiol.* CXXI, 400.
HARTRIDGE, H. (1918). *J. Physiol.* LII, 175.
—— (1923). *J. Physiol.* LVII, 52.
HECHT, S. (1928). *J. gen. Physiol.* XI, 225.
—— (1934). *Handbook of Experimental Psychology*, p. 734. Ed. by C. Murchison.
—— (1935). *J. gen. Physiol.* XVIII, 767.
HELMHOLTZ, H. VON (1866). *Physiological Optics*, Eng. trans, 1924, II, 179, 237–8, 266, 281.
KÖNIG, A. (1897). *S.B. Akad. Wiss. Wien*, p. 559.
KÖNIG, A. AND BRODHUN, E. (1889). *S.B. Akad. Wiss. Wien*, p. 641.
KRAVKOV, S. V. (1938). *Acta ophthal., Kbh.*, XVI. 385.
LASAREFF, P. (1911). *Pflug. Arch. ges. Physiol.* CXLII, 235.
LEGRAND, Y. (1935). *C.R. Acad. Sci., Paris*, CC, 490.
LYTHGOE, R. J. (1932). *Med. Res. Coun. Spec. Rep. Ser.* No. 173.
LYTHGOE, R. J. & TANSLEY, K. (1929). *Proc. roy. Soc.* B, CV, 60.
REEVES, P. (1918). *Astrophys. J.* XLVII, 143.
SCHLAER, S. (1937). *J. gen. Physiol.* XXI, 165.
STEINHARDT, J. (1936). *J. gen. Physiol.* XX, 185.
STILES, W. S. (1929). *Ill. Res. Tech. Paper*, VIII.

(*Manuscript received* 9 *May* 1938)

RELIABILITY OF MENTAL TESTS

BY ROBERT W. B. JACKSON
Colonial Department, University of London Institute of Education

I. *Introduction and statement of the psychological problems considered* (pp. 267–269).
II. *Application of the method of testing statistical hypotheses to the solution of problems* (1) *and* (2) (pp. 270–277).
III. *Estimation of the trial effect* (pp. 277–278).
IV. *Estimation of the sensitivity of the test* (pp. 278–279).
V. *Examples* (pp. 280–283).
VI. *Relationship between the 'reliability' and the 'sensitivity' of a test* (pp. 284–285).
VII. *Sensitivity of a 'whole' test and that of a 'half' test* (pp. 285–286).
VIII. *Summary* (pp. 286–287).
References (p. 287).

I. INTRODUCTION AND STATEMENT OF THE PSYCHOLOGICAL PROBLEMS CONSIDERED

MENTAL tests may be used for many purposes, the diagnosis of mental deficiency, the grading of pupils, and in vocational guidance and selection. Whatever the purpose, the test is assumed to measure some capacity of the individual tested, and we use the score, or some function of the score, obtained by an individual on the test as an estimate of the individual's ability or capacity.

A test may be considered as reliable if it consistently measures whatever capacity it was designed to measure. The correlation between scores obtained by individuals on (1) duplicate forms of the same test given at different times, or (2) two applications of the same test, or (3) two parts of the same test, has been used as an estimate of the reliability, and the correlation coefficient so obtained has been termed the 'reliability coefficient' of the test.[1]

This is the definition, and measure, of reliability in common use. It seems to me, however, that we may consider this problem from another viewpoint. A mental test is, or may be considered as, an instrument for measuring capacity in much the same sense, for example, as a yard-stick is an instrument for measuring length. It may be noticed that the appropriateness of the yard-stick in any given case depends on the relationship between the scale in which it is subdivided (e.g. inches, and tenths of an inch) and the variability of the objects measured. If these differ, for example, by not more than one-tenth of an inch, then, probably, an ordinary yard-stick would not be considered as a sufficiently accurate tool. On the other hand the yard-stick is frequently considered as satisfactory when the objects differ by as much as a foot. This idea seems to be fundamental and I shall use it below, applying it to mental tests. These will be considered as reliable or appropriate when the inaccuracy of measurement by means of these tests is not very great compared with the variability of the capacity measured. The mental test, of course, may be a poor measuring instrument, and we generally wish to know how accurately, or inaccurately, it measures the capacity it was designed to measure. To answer this question we would proceed exactly as before in obtaining the experimental results, but, as will be shown later, the correlation coefficient is not the appropriate measure to use in this case.

---

[1] In the third case the reliability of the whole test is found by using the Spearman-Brown formula.

The score of an individual in a particular trial of a test will be composed of a measure of the capacity of the individual, a measure of the effect common to all individuals on that particular trial, and a measure of numerous other effects including errors of sampling which we cannot measure separately but must group together. We may subdivide the measure of the effect common to all individuals on a particular trial into two parts: (1) a measure of the effect common to all individuals irrespective of the particular trial under consideration and (2) a measure of the trial effect, i.e. a measure of the effect determined by the particular trial under consideration after the common effect (1) has been allowed for. It will be seen that the trial effect must be considered independently of the measurement of the individual's capacity since it may exist whether or not the test actually measures the capacity of the individual. The interpretation which may be placed on reliability depends on the purpose for which the test is to be used. For example, in the diagnosis of mental deficiency we calculate an estimate of the individual's mental age, or intelligence quotient, from the score obtained on the test. In this case we should conclude that a test is satisfactory if it measures the capacity of the individuals accurately, and if the trial effect is not significant; if the trial effect is significant our estimate of the individual's mental age, or intelligence quotient, will depend on the particular trial of the test from which the score is taken. On the other hand, in the grading of pupils we are interested mainly in the position of the individual in the classification, and here the trial effect is unimportant if we may assume that it is independent of the particular individual tested. In general, however, we shall be interested in the actual rather than in the relative scores, and must consider the trial effect independently of the measurement of the capacity of the individual.

The present paper, therefore, is concerned with four separate problems:

(1) To determine if there is a significant trial effect.

(2) To determine if the mental test actually measures the capacity of the individuals tested.

(3) To estimate the trial effect if it exists.

(4) To obtain an estimate of the relative importance of the random errors of measurement, with respect to the true measurement of the capacity of the individual, in determining the score of the individual on the mental test.

In the following pages I have assumed that a linear relation holds between the effects enumerated above and have considered the grouped effects as error; this grouped effect could more properly be termed the residual effect but, as we are unable to distinguish between the pure error and other effects of this group, we may consider the total residual effect as error. It was realized that the relation between these effects may not be linear but, as the true relation is in any case unknown, this form was used because the analysis was simplified and I could apply directly the methods of testing statistical hypotheses, and of estimation, devised and developed by Neyman and Pearson (1, 2, 3).

I have assumed also that the capacity of the individual does not change in the interval elapsing between the first test and the retest. The question of possible changes of capacity has been discussed by Thouless (8) who points out that variations in the experimental results from repeated tests may be due to two factors—unreliability of the test and changes in capacity of the individuals tested. The problem of distinguishing between, and measuring, the effects of these two factors could be attacked by methods similar to the ones outlined in the following pages, but this is an entirely different problem from the one which I have considered. The experimental results which have been analysed below are not affected by changes in capacity as they were obtained by dividing the scores on a single application of the test into two equivalent parts. The results in some cases may be affected, however, and the interpretation of the results obtained by applying the following method would depend on the particular case considered.

II. APPLICATION OF THE METHOD OF TESTING STATISTICAL HYPOTHESES TO THE SOLUTION OF PROBLEMS (1) AND (2)

The event, E, that we observe consists of a set of test scores, in general s test scores, say, for each of n individuals. If correlation methods are used it is necessary to restrict ourselves to the case $s=2$, but the following method is applicable for any value of s greater than unity. The arguments will be developed for the case $s=2$ since the data available have been collected in a form suitable for correlation analysis, as shown in Table I, where y_{st} represents the score of the tth individual on the sth trial of the test. I shall consider the case $s=1, 2$; $t=1, 2, \ldots, n$.

Table I

Individual	Trial I	Trial II
1	y_{11}	y_{21}
2	y_{12}	y_{22}
$\vdots$	$\vdots$	$\vdots$
n	y_{1n}	y_{2n}
Totals	$\sum_t y_{1t}$	$\sum_t y_{2t}$

The data may be obtained in several ways, as indicated in § I, by the use of:

(1) Duplicate forms of the same test; i.e. two separate forms of the test are available, the first form is used and later, after a suitable period of time has elapsed, the second form is used on the same class. It is generally assumed that the two forms of the test are equivalent, in a psychological sense, and that the capacities of the individuals have not changed in the interval elapsing between the trials.

(2) Two applications of the same test; i.e. the test is used once and repeated, after a suitable period of time, on the same class. It is assumed that the capacities of the individuals have not changed in the interval, and that the practice effect, if any, is independent of the particular individual tested.

(3) A division of the scores by some method into two parts, generally by taking the totals of the scores on the odd and even items for each individual. It is assumed that the two parts are equivalent, in a psychological sense, and that the difference in difficulty of the test items on the two parts is not sufficient to invalidate the division.

It is realized(6) that the psychological interpretation of the results obtained by different experimental methods may not be the same in each case, but from the statistical point of view the problem is the same. The concepts introduced are general, their exact interpretation in particular cases is a psychological rather than a statistical problem.

Assume that the score y_{st} of the tth individual on the sth trial of a test may be represented as a sum of independent factors or components. We are interested in an effect common to all individuals on the test irrespective of the particular trial, a trial effect, an individual effect, and the residual, or error, effect.

We may express this in a mathematical form by assuming that

$$y_{st}=A+B_s+C_t+z_{st}, \qquad \text{......(1)}$$

where $s=1, 2$; $t=1, 2, \ldots, n$ represents the score of the tth individual on the sth trial of the test; A is a measure of the effect common to all individuals irrespective of the particular trial under consideration; it will be considered as a constant, and is defined as the arithmetic mean of the effects for all trials; B_s is a measure of the trial effect; it will also be considered as a constant; C_t is a measure of the individual effect, i.e. a measure of some capacity of the individual measured by the test; z_{st} is the error of measuring by means of the test applied.

We assume, further, that

(1) z_{st} is independent of C_t,

(2) z_{st} is normally distributed about zero with constant standard deviation, σ, say, constant for all trials and all individuals, and

(3) in the population from which we are sampling, C_t is normally distributed about zero with the standard deviation, σ_c, independent of the trial.

The tests are generally so constructed that the individual with the least capacity can answer the easiest questions, while the individual with the greatest capacity will be unable to complete the test. Because of this type of construction the distribution of test scores is generally such that it can be adequately represented by a normal distribution, and hence the above assumptions will define a mathematical model which should adequately represent the observed events.

Finally, since we have considered A as a common effect and defined it as the arithmetic mean of the effects for all trials and all individuals, it is necessary that

$$\sum_s B_s=0. \qquad \text{......(2)}$$

In the particular case considered, (2) becomes

$$\left.\begin{aligned} &B_1+B_2=0 \\ \text{or} \quad &B_1=-B_2=-B, \text{ say,}\end{aligned}\right\} \qquad \text{......(3)}$$

where B is now the measure of the trial effect, and we may write (1) in the form

$$\left.\begin{aligned} y_{1t}&=A-B+C_t+z_{1t},\\ y_{2t}&=A+B+C_t+z_{2t}.\end{aligned}\right\} \qquad \ldots\ldots(4)$$

Therefore, the score of the tth individual on the sth trial of the test is assumed to be composed of a constant term $A \pm B$, plus a measure of the capacity of the individual, C_t, and a random term z_{st}. If σ is large in comparison with σ_c, it is easily seen that we may frequently get a deviation from the true capacity, C_t, as large or larger than σ_c units due to chance alone if we use the score on the test as an estimate of the capacity of the individual. On the other hand, if σ is small in comparison with σ_c, then we would seldom make an error by chance alone as great or greater than σ_c units in using the score as an estimate of the capacity of the individual tested. This effect is, of course, quite different from, and independent of, the bias introduced by the trial effect as considered above.

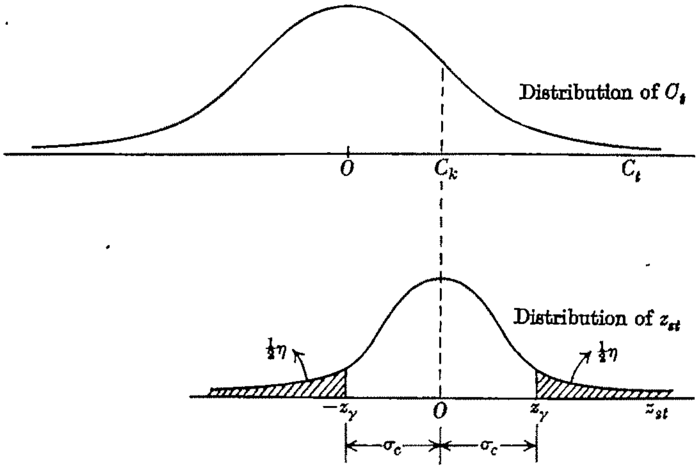


Fig. 1. Illustrating the possible effect of the z_{st} term in the score y_{st} of an individual.

The chance deviations may be very great in any particular case and we cannot control these, but for a particular value of γ, where γ is defined as σ_c/σ, we can determine the probability η, say, of making an error greater than or equal to σ_c units due to chance alone in using the score as an estimate of the capacity of the individual. I will describe γ as the sensitivity[1] of the mental test. It will be seen that γ is independent of the scale used in marking the test. The position is illustrated in Fig. 1.

---

[1] The relationship between the 'reliability' and the 'sensitivity' of a test is discussed below, see p. 284.

It is seen that the problem reduces to one of finding a value of η such that

$$P\{-z_\gamma \leqslant z_{st} \leqslant z_\gamma\} = 1 - \eta, \qquad \ldots\ldots(5)$$

where $z_\gamma = \sigma_c$. Since z_{st} is assumed to be normally distributed about zero with constant standard deviation, σ, this reduces to finding η such that

$$\frac{1}{\sqrt{(2\pi)}} \int_{-z_\gamma/\sigma}^{z_\gamma/\sigma} e^{-\frac{1}{2}x^2}\, dx = 1 - \eta, \qquad \ldots\ldots(6)$$

But $z_\gamma/\sigma = \sigma_c/\sigma = \gamma$, and hence we may obtain η directly from the tables of the normal integral for any γ. And, conversely, if we have a particular value of η, say one that we have chosen in advance, we can obtain γ from the tables of the normal integral. For example, if $\gamma = 1{\cdot}96$, then $z_\gamma/\sigma = 1{\cdot}96$, and from the tables we find $\eta = 0{\cdot}05$. This means that in using the scores as estimates of capacity we would expect to make an error as great or greater than σ_c units by chance alone only once in twenty times. The arguments are similar for any other value of γ or η.

In the first problem we wish to determine if there is a significant trial effect. This is equivalent to testing the hypothesis that the trial effect is zero; the corresponding statistical hypothesis, H_1, may be written in the form

$$H_1 : B = 0. \qquad \ldots\ldots(7)$$

In the second problem we wish to determine if the mental test actually measures the capacity of the individuals tested. This is equivalent to testing the hypothesis that the sensitivity of the test is zero; the corresponding statistical hypothesis, H_2, may be written in the form

$$H_2 : \gamma = 0. \qquad \ldots\ldots(8)$$

The mathematical development of the tests to be used in testing these statistical hypotheses, and of the method of constructing the confidence intervals to be used in estimating B or γ if these hypotheses are rejected, will be published later (9). The tests to be used, and the method of constructing the confidence intervals, will be stated here and their use explained by means of examples. This division has been made because it was thought that practical psychologists and educationists would be interested mainly in the method and results, but not so much in the mathematical theory underlying these. On the other hand this theory may be of more general interest. There are probably many other examples in psychology, and possibly in other fields, where the same, or a similar, general method could be used. The next few paragraphs will be devoted to a brief description of the theory of testing statistical hypotheses and of estimation introduced and developed by Neyman and Pearson(1, 2, 3).

Denote by W the same space, the co-ordinates of which are the possible results of experiment, and by w any region in this sample space. For example, in the case of mental tests the scores would be considered as the co-ordinates defining the sample space, and the observed scores on any particular experiment would determine a point, E, called the sample point, in the space of all such possible results. Denote by $P\{E \epsilon w\}$ the probability that E, as determined by some experimental results, will fall within the region w, and by $P\{E \epsilon w/H\}$ the probability of E falling within w as determined by some hypothesis, H.

Any test of a statistical hypothesis, H_0, reduces to a rule of rejecting H_0 when the sample point, E, falls within some specified region, w_0, called critical, and of refraining from doing so (i.e. 'accepting H_0') in other cases. The choice of critical regions is made having regard to probabilities of errors which may arise in testing hypotheses. There are two kinds of errors: (1) we may reject the hypothesis tested, H_0, when it is true. The probability of the first kind of error is thus the probability, as determined by H_0, of E falling within the critical region, $P\{E \epsilon w_0/H_0\}$. (2) We may fail to reject H_0 when, in fact, it is wrong, and therefore when some other hypothesis, H' say, contradictory to H_0 is correct. If A_0 denote the region $W - w_0$, then the probability of the second kind of error will be $P\{E \epsilon A_0/H'\}$. The region A_0 is called the region of acceptance. The critical regions used in the tests of the statistical hypotheses in the following sections are the best possible in the sense that (1) the probability of errors of the first kind is controlled at a fixed level, and (2) the probability of errors of the second kind is reduced to as low a level as possible for the class of hypotheses in which we are interested. It will be noticed that the tests of these hypotheses are of the form of the tests used in the Analysis of Variance (R. A. Fisher(5)) method of analysing experimental results.

The method of estimating an unknown parameter θ, say (e.g. B or γ in our case), consists essentially in determining two functions $\underline{\theta}(E)$ and $\overline{\theta}(E)$, having the property that the probability of $\underline{\theta}(E)$ falling short of the true value of θ, say θ_0, and at the same time of $\overline{\theta}(E)$ exceeding θ_0 is exactly equal to α, a number chosen in advance and as close to unity as desired. The number α is called the 'confidence coefficient', and the interval extending from $\underline{\theta}(E)$ to $\overline{\theta}(E)$ is called the 'confidence interval'. If the functions $\underline{\theta}(E)$ and $\overline{\theta}(E)$ are determined then the process of estimation of θ consists in (1) obtaining the experimental results determining the sample point, E, (2) calculating the values of $\underline{\theta}\,(E)$ and $\bar{\theta}\,(E)$ and (3) stating then that $\underline{\theta}\,(E) \leqslant \theta \leqslant \bar{\theta}\,(E)$. The justification for this procedure is that the probability of the statement being correct is equal to α—the value chosen in advance.

Neyman has shown that in general there is an infinity of functions $\underline{\theta}\,(E)$ and $\bar{\theta}\,(E)$ satisfying the above conditions, and therefore an infinity of systems of confidence intervals (3). The particular system of confidence intervals used below possesses the following properties: (1) the confidence intervals cover the true value of θ, say θ_0, in the proportion α of all such possible cases, and (2) if θ' is not the true value of θ, the confidence intervals cover θ' as rarely as possible. (The system used is known as a 'short unbiassed' system of confidence intervals).

Table II

	Trial I	Trial II	Totals	Squares of totals	Differences	Squares of differences
Individual	y_{1t}	y_{2t}	$y_{1t}+y_{2t}$	$(y_{1t}+y_{2t})^2$	$y_{2t}-y_{1t}$	$(y_{2t}-y_{1t})^2$
1	y_{11}	y_{21}				
2	y_{12}	y_{22}				
$\vdots$	$\vdots$	$\vdots$				
n	y_{1n}	y_{2n}				
Totals	$\sum_t y_{1t}$	$\sum_t y_{2t}$	$\sum_t (y_{1t}+y_{2t})$	$\sum_t (y_{1t}+y_{2t})^2$	$\sum_t (y_{2t}-y_{1t})$	$\sum_t (y_{2t}-y_{1t})^2$

Table III

Statistic	Definition
$\bar{v}$	$\frac{1}{n}\sum_t (y_{2t}-y_{1t})$
S_u^2	$\frac{1}{n}\left\{\sum_t (y_{1t}+y_{2t})^2 - \frac{(\sum_t y_{1t}+\sum_t y_{2t})^2}{n}\right\}$
S_v^2	$\frac{1}{n}\left\{\sum_t (y_{2t}-y_{1t})^2 - \frac{(\sum_t y_{2t}-\sum_t y_{1t})^2}{n}\right\}$
χ_v^2	$\frac{1}{n}\sum_t (y_{2t}-y_{1t})^2$

The statistics to be used in testing the hypotheses H_1 and H_2, and later in the estimation of the trial effect and the sensitivity of the test if we reject these hypotheses, may be calculated as shown in Table II. From the totals of Table II we calculate those in Table III. The calculations may be checked by use of the following identity:

$$\sum_t y_{1t}^2 + \sum_t y_{2t}^2 - \frac{1}{2n}\{\sum_t y_{1t} + \sum_t y_{2t})^2 \equiv \tfrac{1}{2}n\,\{(\bar{v})^2 + S_u^2 + S_v^2\}. \quad \ldots\ldots(9)$$

(A) *Test of the hypothesis* $H_1 : B=0$

(i) From Table III calculate

$$z_1=\tfrac{1}{2}\log_e\left\{\frac{(\bar{v})^2}{S_v^2/(n-1)}\right\} \qquad \text{......(10)}$$

and refer to Fisher's tables of z for $f_1=1$, and $f_2=n-1$,

(ii) reject the hypothesis H_1 if z_1 is greater than the 5% (or 1%) point given in the tables,

(iii) accept the hypothesis H_1 if z_1 is less than the 5% (or 11%) point given in the tables.

(B) *Test of the hypothesis* $H_2 : \gamma=0$

(i) From Table III calculate

$$z_2=\tfrac{1}{2}\log_e\left\{\frac{S_u^2}{S_v^2}\right\}, \qquad \text{......(11)}$$

and refer to Fisher's tables of z for $f_1=f_2=n-1$,

(ii) reject the hypothesis H_2 if z_2 is greater than the 5% (or 1%) point given in the tables,

(iii) accept the hypothesis H_2 if z_2 is less than the 5% (or 1%) point given in the tables.

If we know from past experience, and from the test of H_1 for the observed data, that we consider that the trial effect is zero then we may test the hypothesis H_2 under these conditions. The test is, of course, different from that considered above (B), and should be used only when the two conditions are satisfied.

(C) *Test of the hypothesis* $H_2 : \gamma=0$, *when* B *is known to be zero*

(i) From Table III calculate

$$z_2'=\tfrac{1}{2}\log_e\left\{\frac{S_u^2/(n-1)}{\chi_v^2/n}\right\} \qquad \text{......(12)}$$

and refer to Fisher's tables of z for $f_1=n-1$, and $f_2=n$,

(ii) reject the hypothesis H_2 if z_2' is greater than the 5% (or 1%) point given in the tables,

(iii) accept the hypothesis H_2 if z_2' is less than the 5% (or 1%) point given in the tables.

The question of whether we should use the 5% or 1% point is a personal one, and depends upon what risk of rejecting the hypothesis tested when it is true (i.e. the probability of the first kind of errors) we consider as negligible or permissible. The general practice seems to be to reject the hypothesis tested if the value calculated from the sample is above the 1% point, to remain in doubt if it is between the 5% and 1% points, and to accept the hypothesis tested if it is below the 5% point.

III. ESTIMATION OF THE TRIAL EFFECT

If we accept the hypothesis $H_1 : B=0$, then we may conclude that the trial effect is not significant. However, if we reject H_1 then we conclude that the trial effect is significant, and the problem arises of estimating this trial effect, B.

As shown above, this problem reduces to that of finding the upper and lower bounds of the confidence interval, $\underline{B}$ and $\overline{B}$ respectively, such that the probability of the statement

$$\underline{B} \leqslant B \leqslant \overline{B} \qquad \ldots\ldots(13)$$

being correct is equal to $1-\epsilon$, where ϵ is any small number chosen in advance. In making the statement concerning B in the form (13) we know that we shall be correct in the proportion $1-\epsilon$ of all cases. We find that

$$\underline{B}=\frac{1}{2}\left(\bar{v}-S_v\sqrt{\frac{l}{1-l}}\right),$$

$$\overline{B}=\frac{1}{2}\left(\bar{v}+S_v\sqrt{\frac{l}{1-l}}\right), \qquad \ldots\ldots(14)$$

where l is chosen such that

$$\frac{\Gamma_{\frac{1}{2}n}}{\sqrt{\pi}\,\Gamma_{\frac{1}{2}(n-1)}}\int_0^l z^{-\frac{1}{2}}(1-z)^{\frac{1}{2}(n-3)}\,dz=1-\epsilon. \qquad \ldots\ldots(15)$$

We may obtain l by interpolating in Pearson's Tables of the Incomplete Beta-Function for any n and ϵ (4). The choice of a particular value of ϵ to use is a personal one, but generally $\epsilon=0{\cdot}05$, or $\epsilon=0{\cdot}02$, is used.

If n is larger than 101 then these tables could not be used and the first column of a small table ((7) Table V, p. 92) published by Johnson and Neyman could be recommended. This gives directly the values of $1-l$, which the authors denote by z_α, corresponding to several large values of n and to $\alpha=0{\cdot}05$ and $\alpha=0{\cdot}01$. The table is so arranged as to allow an easy interpolation for values of n not included in the table.

If $n \leqslant 31$ the alternative method of calculating $\underline{B}$ and $\overline{B}$ consists in using Fisher's tables of t(5). The equations for $\underline{B}$ and $\overline{B}$ are

$$\left.\begin{aligned}\underline{B}&=\frac{1}{2}\left(\bar{v}-\frac{S_v}{\sqrt{(n-1)}}\,t_\epsilon\right),\\ \overline{B}&=\frac{1}{2}\left(\bar{v}+\frac{S_v}{\sqrt{(n-1)}}\,t_\epsilon\right),\end{aligned}\right\}\qquad\text{......(16)}$$

where t_ϵ is found by entering this table with degrees of freedom equal to $n-1$, and $P=\epsilon$.

IV. ESTIMATION OF THE SENSITIVITY OF THE TEST

If we accept the hypothesis $H_2: \gamma=0$, then we may conclude that the mental test is not sensitive enough to measure the capacities of the individuals tested and hence it would not be a satisfactory test. If we reject H_2 we conclude that the test is sensitive enough to measure the capacities of the individuals tested. How sensitive a test must be before it can be considered as satisfactory is, of course, a personal question, and could only be answered objectively by analysing the scores from some 'standard' test which has been generally accepted as giving 'satisfactory' results. This question of fixing a lower limit for the sensitivity is not a purely statistical problem. However, if we reject H_2 we would generally wish to know how sensitive the test is, or, in other words, to estimate γ, the sensitivity of the test.

As in the case of the test of the hypothesis H_2, this problem must be divided into two parts:

(A) Estimation of γ if the value of B is unknown, and

(B) Estimation of γ if the value of B is known to be zero.

We are justified in using (B) only if we know from past experience, and from the test of the hypothesis H_1 for the observed data, that the trial effect is not significant. Otherwise case (A) is the appropriate one to use.

(A) *Estimation of γ if the value of B is unknown*

The confidence interval to be used is denoted by

$$\underline{\gamma}=+\sqrt{\frac{\underline{\theta}-1}{2}}\leqslant\gamma\leqslant+\sqrt{\frac{\overline{\theta}-1}{2}}=\overline{\gamma},\qquad\text{......(17)}$$

where

$$\left.\begin{aligned}\underline{\theta}&=\frac{S_u^2}{S_v^2}\left(\frac{k_1}{1-k_1}\right),\\ \overline{\theta}&=\frac{S_u^2}{S_v^2}\left(\frac{1-k_1}{k_1}\right),\end{aligned}\right\}\qquad\text{......(18)}$$

and k_1 is found from the equation

$$\frac{\Gamma_{(n-1)}}{(\Gamma_{\frac{1}{2}(n-1)})^2}\int_0^{k_1} z^{\frac{1}{2}(n-3)}(1-z)^{\frac{1}{2}(n-3)}\,dz=\tfrac{1}{2}\epsilon \qquad \text{......(19)}$$

and the tables of the Incomplete Beta-Function. Or, alternatively, by using the relation

$$\tfrac{1}{2}\log_e\left(\frac{1-k_1}{k_1}\right)=z_{\frac{1}{2}\epsilon} \qquad \text{......(20)}$$

finding $z_{\frac{1}{2}\epsilon}$ from Fisher's tables of z for $f_1=f_2=n-1$, and solving (20) for k_1. For values of $n>101$, k_1 can be determined only from Fisher's tables. The property of the confidence interval is again that the probability of the statement in the form $\underline{\gamma}\leqslant\gamma\leqslant\bar{\gamma}$ being correct is equal to the number $1-\epsilon$ fixed in advance.

(B) *Estimation of γ is B is known to be zero*

The confidence interval is defined as in (17), but $\underline{\theta}$ and $\bar{\theta}$ are now defined by

$$\left.\begin{aligned}\underline{\theta}&=\left(\frac{1-m_2}{m_2}\right)\frac{S_u{}^2}{\chi_v{}^2},\\ \bar{\theta}&=\left(\frac{1-m_1}{m_1}\right)\frac{S_u{}^2}{\chi_v{}^2},\end{aligned}\right\} \qquad \text{......(21)}$$

where m_1 and m_2 are defined by the equations

$$(m_2)^{\frac{1}{2}(n-1)}(1-m_2)^{\frac{1}{2}n}=(m_1)^{\frac{1}{2}(n-1)}(1-m_1)^{\frac{1}{2}n}, \qquad \text{......(22)}$$

$$\frac{\Gamma_{\frac{1}{2}(2n-1)}}{\Gamma_{\frac{1}{2}n}\Gamma_{\frac{1}{2}(n-1)}}\int_{m_1}^{m_2} y^{\frac{1}{2}(n-3)}(1-y)^{\frac{1}{2}(n-2)}\,dy=1-\epsilon. \qquad \text{......(23)}$$

For values of $n>20$ we may determine m_1 and m_2 from the equation

$$\begin{aligned}&\frac{\Gamma_{\frac{1}{2}(2n-1)}}{\Gamma_{\frac{1}{2}n}\Gamma_{\frac{1}{2}(n-1)}}\int_0^{m_1} y^{\frac{1}{2}(n-3)}(1-y)^{\frac{1}{2}(n-2)}\,dy\\ &\qquad=\frac{\Gamma_{\frac{1}{2}(2n-1)}}{\Gamma_{\frac{1}{2}n}\Gamma_{\frac{1}{2}(n-1)}}\int_{m_2}^{1} y^{\frac{1}{2}(n-3)}(1-y)^{\frac{1}{2}(n-2)}\,dy=\tfrac{1}{2}\epsilon. \qquad \text{......(24)}\end{aligned}$$

Or, alternatively, for values of $n>20$ we may obtain the values of m_1 and m_2 by using the following relations and Fisher's tables of z:

$$\tfrac{1}{2}\log_e\left\{\left(\frac{n}{n-1}\right)\left(\frac{m_2}{1-m_2}\right)\right\}=z_{\frac{1}{2}\epsilon}, \qquad \text{......(25)}$$

and

$$\tfrac{1}{2}\log_e\left\{\left(\frac{n-1}{n}\right)\left(\frac{1-m_1}{m_1}\right)\right\}=z'_{\frac{1}{2}\epsilon}, \qquad \text{......(26)}$$

where $z_{\frac{1}{2}\epsilon}$ is obtained from the tables for $f_1=n-1$ and $f_2=n$, and $z'_{\frac{1}{2}\epsilon}$ from the same tables for $f_1=n$ and $f_2=n-1$. For values of $n>100$ we can obtain these values only from Fisher's tables of z.

V. EXAMPLES

The data for the following examples have been very kindly supplied by Dr Palmer O. Johnson of the University of Minnesota, U.S.A. They refer to a set of Comprehensive Examinations in Biological Sciences, prepared by Dr Johnson, and are the results for a class of 132 students at the University of Minnesota in 1935. I wish to take this opportunity to thank him for so kindly supplying these data. The original test scores are given in Table IV; it will be seen that the method used to obtain the experimental results for testing reliability is the third of those considered in Sections (1) and (2), i.e. by taking the totals of the scores on the odd and even test items for each individual. The set of tests is given in three parts: Part I, Acquisition of Vocabulary, Part II, Acquisition of Principles, and Part III, Application of Principles in Problems, or, more fully, The Ability to Apply Principles in the Solution of Biological Problems.

Table IV. *Original test scores*

	Part I		Part II		Part III			Part I		Part II		Part III	
Ind.	Odd	Even	Odd	Even	Odd	Even	Ind.	Odd	Even	Odd	Even	Odd	Even
1	46	52	63	57	34	35	32	43	39	73	64	41	40
2	38	38	71	67	42	36	33	58	54	59	67	43	37
3	64	63	70	74	41	44	34	54	50	91	97	48	54
4	73	65	119	114	50	52	35	59	59	80	83	52	59
5	61	58	109	111	57	57	36	68	65	105	126	59	62
6	34	33	75	80	30	41	37	51	57	97	96	48	53
7	57	49	88	87	33	36	38	44	37	77	72	39	36
8	66	63	83	100	55	49	39	77	76	124	116	52	57
9	25	24	68	56	20	14	40	46	57	68	70	34	41
10	30	26	59	57	35	28	41	56	53	101	98	49	45
11	45	33	55	60	43	47	42	44	46	81	67	34	32
12	73	71	106	113	47	56	43	54	46	69	65	44	43
13	45	48	56	56	35	45	44	42	51	73	81	40	44
14	55	63	81	79	51	59	45	28	30	63	74	34	29
15	66	70	102	110	48	51	46	38	39	83	89	44	42
16	49	46	94	95	43	44	47	56	52	80	80	52	54
17	64	65	97	95	40	47	48	50	47	89	77	49	50
18	45	46	84	86	39	39	49	62	54	98	120	44	47
19	61	62	91	100	51	51	50	33	42	73	76	35	39
20	52	46	85	79	41	36	51	27	28	65	71	30	32
21	67	68	106	124	49	56	52	40	37	62	77	30	30
22	59	53	86	107	49	51	53	74	73	114	113	54	61
23	55	55	104	98	41	35	54	64	54	105	96	39	46
24	51	52	78	83	40	35	55	48	51	88	81	35	46
25	50	48	91	99	51	41	56	53	46	78	84	49	52
26	52	53	82	85	43	41	57	56	52	69	71	51	53
27	28	26	64	69	34	41	58	35	41	67	70	36	46
28	42	44	55	55	38	38	59	55	45	79	80	29	41
29	59	55	87	99	40	54	60	50	46	80	95	38	41
30	24	23	50	48	30	22	61	29	28	47	74	36	34
31	42	44	75	103	46	50	62	55	56	68	65	42	49

Table IV (*continued*)

	Part I		Part II		Part III			Part I		Part II		Part III	
Ind.	Odd	Even	Odd	Even	Odd	Even	Ind.	Odd	Even	Odd	Even	Odd	Even
63	54	61	93	109	44	54	98	47	51	83	93	40	46
64	52	46	78	89	37	45	99	49	47	98	87	37	45
65	44	43	51	45	34	30	100	30	32	61	65	18	28
66	62	56	92	101	46	46	101	45	50	80	73	39	46
67	55	56	76	80	36	36	102	67	63	70	77	40	46
68	56	62	105	102	57	60	103	15	28	60	65	30	30
69	49	47	55	64	32	39	104	30	28	66	56	34	39
70	56	61	86	114	50	56	105	44	41	71	75	31	34
71	37	45	71	56	30	19	106	45	68	85	85	46	39
72	49	44	70	68	31	43	107	24	21	43	55	26	30
73	37	37	68	84	28	23	108	51	54	65	56	32	32
74	41	42	81	88	39	35	109	42	42	53	50	35	37
75	47	47	81	77	48	47	110	47	45	88	92	45	47
76	55	51	65	71	39	45	111	38	43	68	63	41	42
77	59	62	104	106	49	47	112	56	49	93	96	46	49
78	53	49	88	95	43	43	113	50	52	91	89	47	50
79	41	42	78	77	32	38	114	63	66	101	114	56	54
80	48	52	84	87	40	41	115	45	39	94	103	40	51
81	62	51	92	82	47	49	116	56	53	91	103	41	41
82	41	40	84	89	35	46	117	57	55	73	83	33	35
83	50	40	78	89	48	52	118	64	66	99	98	47	48
84	40	47	66	74	25	33	119	59	67	99	113	45	54
85	68	64	94	100	53	60	120	35	29	66	59	40	43
86	43	45	52	51	39	35	121	53	41	78	73	40	36
87	49	47	61	76	38	36	122	44	52	56	54	37	44
88	65	62	96	97	43	49	123	54	57	93	86	48	49
89	50	51	53	60	27	39	124	48	47	85	65	38	33
90	31	30	49	70	22	34	125	47	36	58	61	36	38
91	55	51	90	98	49	35	126	52	45	92	101	43	43
92	45	40	69	65	31	33	127	45	42	75	65	31	33
93	40	46	52	51	37	37	128	23	25	66	75	27	33
94	38	42	40	54	41	31	129	43	43	69	75	44	36
95	55	55	82	76	40	45	130	73	66	111	122	50	55
96	43	44	90	80	37	33	131	32	29	73	64	35	31
97	46	56	108	113	54	60	132	44	45	73	70	41	42

I analysed these data as suggested in Tables II and III, and obtained the results shown in Table V.

Table V

	Test		
Statistic	Part I	Part II	Part III
$\bar{v}$	−0·5758	2·9621	2·0682
S_u^2	522·9984	1211·4126	272·6996
S_v^2	28·6685	81·5818	30·3514
χ_v^2	29·0000	—	—

Test of the hypothesis $H_1 : B = 0$. Substituting the values of $\bar{v}$ and S_v^2 from Table V in equation (10) we obtain the results shown in Table VI, where z_1 (5%) and z_1 (1%) are the 5 and 1% points of the distribution of z, respectively, obtained from Fisher's tables of z for $f_1 = 1$, and

$$f_2 = n - 1 = 131.$$

As will be seen from the last column of Table VI, the trial effect may be considered to be zero for Part I, but is different from zero for Parts II and III.

Table VI

Test	z_1	z_1 (5%)	z_1 (1%)	Result
Part I	0·2076	0·6821	0·9609	Accept H_1
Part II	1·3227	0·6821	0·9609	Reject H_1
Part III	1·4578	0·6821	0·9609	Reject H_1

Test of the hypothesis $H_2: \theta=1$, *or* $\gamma=0$. Substituting the values of S_u^2 and S_v^2 from Table V in equation (11) we obtain the results shown in Table VII, where z_2 (5%) and z_2 (1%) are the 5 and 1% points of the distribution of z, respectively, obtained from Fisher's tables of z for $f_1=f_2=n-1=131$. As will be seen from the last column of Table VII, we should reject H_2 in each case, and hence we should conclude that these tests are sensitive enough to measure the capacities of the individuals tested.

Table VII

Test	z_2	z_2 (5%)	z_2 (1%)	Result
Part I	1·4519	0·1428	0·2025	Reject H_2
Part II	1·3490	0·1428	0·2025	...
Part III	1·0978	0·1428	0·2025	...

It may be argued that I should have tested H_2 for Part I by the test given under (C) above, but I am not certain that the trial effect will always be zero for this test, and in any case I would have reached the same conclusion and rejected H_2.

Estimation of the trial effect. Since I have accepted H_1 for Part I the question of estimating B, the trial effect, does not arise for this case. The confidence intervals for B for the Part II and Part III tests are given in Table VIII. I used the confidence coefficient $1-\epsilon=0\cdot98$.

Table VIII

Test	Confidence interval for B
Part II	$0\cdot549 \leqslant B \leqslant 2\cdot414$
Part III	$0\cdot467 \leqslant B \leqslant 1\cdot601$

Estimation of γ, *the sensitivity of the test.* Since I am not certain that B will always be zero for Part I, I have used case (A) (p. 278) for all these tests. The confidence coefficient used is $1-\epsilon=0\cdot98$, and the values were calculated using Fisher's tables of z. The results are shown in Table IX.

These values of γ are rather small; it would seem that the random errors of measurement, as I have termed the residual effect, may be very important in determining the actual score of an individual on the test.

Table IX

Test	Confidence interval for γ
Part I	$2{\cdot}36 \leqslant \gamma \leqslant 3{\cdot}63$
Part II	$2{\cdot}11 \leqslant \gamma \leqslant 3{\cdot}26$
Part III	$1{\cdot}58 \leqslant \gamma \leqslant 2{\cdot}50$

I calculated the correlation coefficient for Part I and found $r = 0{\cdot}90$; this test would thus be considered as satisfactory if we used the usual methods. It seems probable, therefore, that many of the mental tests in common use are not satisfactory instruments for measuring capacity. However, the value of S_u^2/S_v^2 does not depend on the value of r alone, although if the assumptions which I have made are satisfied S_u^2/S_v^2 will be of the order of $(1+r)/(1-r)$, and it would be necessary to calculate the confidence intervals for γ from the formulae given above, rather than to attempt to estimate them from the value of r.

I have assumed that the test scores are normally distributed with the same constant standard deviation on each trial of a particular test. We may test the hypothesis that the test scores are normally distributed on each trial by calculating the estimates of β_1 and β_2, and testing whether these differ significantly from zero and three respectively (see *Tables for Statisticians and Biometricians*, Part II, Table XXXVII *bis*). The hypothesis that the standard deviation is the same for each trial of a test may be tested by calculating

$$z = \tfrac{1}{2} \log_\epsilon \left\{ \frac{S^2 y_{1t}}{S^2 y_{2t}} \right\}; f_1 = f_2 = n - 1, \qquad \ldots\ldots(27)$$

and rejecting the hypothesis if z is significantly greater than zero (Fisher's z-test). The results for the three mental tests are given in Table X. We may accept the hypothesis that the test scores are normally distributed in each trial for five of the six cases, the sixth case—for the even scores on Part II—is doubtful. We may accept the hypothesis that the standard deviations are equal for the two trials in all three cases. These results would indicate, in this case at least, that the mathematical model which I have used adequately represents the observed events.

Table X. *Tests of the hypotheses* (1) $\beta_1 = 0$, (2) $\beta_2 = 3$ *and* (3) $\sigma_{\text{odd}} = \sigma_{\text{even}}$

Test	Trial	β_1	β_1 (5%)	β_1 (1%)	β_2	β_2 (5%)	β_2 (1%)	z	z (5%)
Part I	Odd	0·032	0·117	0·246	2·95	2·41	2·25	0·024	0·143
	Even	0·019			2·72				
Part II	Odd	0·009	0·117	0·246	2 55	2·41	2·25	0·104	0·143
	Even	0·059			2·27				
Part III	Odd	0·015	0·117	0·246	2·74	2·41	2·25	0·134	0·143
	Even	0·023			2·87				

VI. RELATIONSHIP BETWEEN THE 'RELIABILITY' AND THE 'SENSITIVITY' OF A TEST

It follows from equations (4) that

$$\rho = \frac{\gamma^2}{1+\gamma^2}, \qquad \ldots\ldots(28)$$

where ρ is the population value of the correlation coefficient between y_{1t} and y_{2t} or, what is usually called the reliability of the test. Table XI gives the values of γ corresponding to certain values of ρ. It is seen that for $\rho = 0{\cdot}94$, σ is roughly $\sigma_0/4$, so that even for this large value of ρ we should conclude that the random errors of measurement may be very important in determining the actual score of the individual on the test.

Table XI

ρ	0·0	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
γ	0·00	0·33	0·50	0·66	0·82	1·00	1·22	1·53	2·00	3·00
ρ	0·91	0·92	0·93	0·94	0·95	0·96	0·97	0·98	0·99	1·00
γ	3·18	3·39	3·64	3·96	4·36	4·90	5·69	7·00	9·95	∞

The use of γ, rather than ρ, as a measure of the reliability of the test has several advantages. I introduced the term 'sensitivity' of the test not only because the concept of sensitivity seemed to me to be of fundamental importance, but also in order to emphasize the desirability of using γ. First, from a mathematical point of view ρ and γ are equivalent, i.e. ρ is a function of γ; but this is not the problem. The true value of γ, or ρ, is in any case unknown; our problem is to estimate the value of γ. It may be argued that we could estimate ρ, and then γ, but this introduces an unnecessary step and, what is more important, it is not the best method to use in estimating γ. The error of estimation of ρ would appear also in our estimate of γ, and the effect of this error on our estimate of γ is difficult to determine. The method of estimating γ given above is the simplest and the best that can be used.

Secondly, it is seen from Table XI that, for values of $\rho > 0{\cdot}9$, a small change in the value of ρ corresponds to a large change in the value of γ. For example, a test for which $\rho = 0{\cdot}99$ is more than twice as sensitive as a test for which $\rho = 0{\cdot}95$. Thus it may happen that a test is actually much more sensitive than another but the difference in reliability might appear to be insignificant if our judgment is based on the values of the correlation coefficients.

Thirdly, the probability distributions of the statistics used in the above method are known, and are of a particularly simple form. The probability distribution of r is known, of course, but it is difficult to evaluate the integral of the probability distribution in the cases for which $\rho \neq 0$—i.e. the cases in which we should be interested if we wished to estimate ρ.

Finally, the methods outlined above could be generalized for values of $s > 2$ (i.e. in cases of more than two trials), but the correlation method is applicable only to the case $s = 2$. It is not necessary to use more than two trials, but it would probably be useful to do so, because we could then study the effect of other factors. For example, it would be interesting to consider the problem of possible changes in the variability of the students in repeated trials, i.e. to develop a test of the hypothesis that a common σ_c exists for all the trials. This problem is outside the scope of the present paper, but the same general method could be used in its solution.

VII. SENSITIVITY OF A 'WHOLE' TEST AND THAT OF A 'HALF' TEST

In the examples given above, the scores on Trial I and Trial II were obtained from the totals of the scores on the odd and even items of the test for each individual. Hence these are actually the scores on halves of the test, and the estimate of sensitivity obtained above is the estimate of the sensitivity of the half-test.

Denote by $\rho_{\frac{1}{2}}$ the correlation, in the population from which we are sampling, between the totals of the scores on the odd and even items of the test, and by ρ_w the correlation, in the population from which we are sampling, between the scores on two comparable forms of the same test. Further, denote by $\gamma_{\frac{1}{2}}$, the sensitivity of the half-test, and by γ_w the sensitivity of the whole test.

From formula (28) we see that

$$\rho_{\frac{1}{2}} = \frac{\gamma_{\frac{1}{2}}^2}{1+\gamma_{\frac{1}{2}}^2} \qquad \text{......(29)}$$

and

$$\rho_w = \frac{\gamma_w^2}{1+\gamma_w^2}. \qquad \text{......(30)}$$

If the two forms of the test are comparable then it follows from the Spearman-Brown formula that

$$\rho_w = \frac{2\rho_{\frac{1}{2}}}{1+\rho_{\frac{1}{2}}}. \qquad \text{......(31)}$$

From (30) it follows that

$$1+2\gamma_w^2 = \frac{1+\rho_w}{1-\rho_w}. \qquad \text{......(32)}$$

Substituting the value of ρ_w from (31) in (32) we have

$$1+2\gamma_w^2=\frac{1+3\rho_{\frac{1}{2}}}{1-\rho_{\frac{1}{2}}}. \qquad \text{......(33)}$$

Substituting the value of $\rho_{\frac{1}{2}}$ from (29) in (33) we have

$$1+2\gamma_w^2=1+4\gamma_{\frac{1}{2}}^2, \qquad \text{......(34)}$$

and we find that

$$\gamma_w=\sqrt{2}\,\gamma_{\frac{1}{2}}. \qquad \text{......(35)}$$

Hence to obtain the estimate of the sensitivity for the whole test from the results of experiments such as those considered above, it would be necessary to multiply the estimate of the sensitivity of the half-test by the factor $\sqrt{2}$. It must be remembered, however, that the above arguments are valid only if the two forms of the test are comparable—a condition which would be rarely satisfied in practice. For this reason the estimates of the sensitivity of the whole test calculated from the results obtained by the use of methods (1) or (2), and the estimate of the sensitivity for the half-test calculated from the results obtained by the use of method (3), will not satisfy equation (35) exactly.

Using equation (35), and the results from Table IX, we find the confidence intervals for γ_w which are shown in Table XII. It is seen that even these 'corrected' estimates of the sensitivity of the tests are rather small, and we should conclude, as before, that the random errors of measurement may be very important in determining the actual score of an individual on the tests.

Table XII

Test	Confidence interval for γ_w
Part I	$3\cdot34\leqslant\gamma_w\leqslant5\cdot13$
Part II	$2\cdot98\leqslant\gamma_w\leqslant4\cdot61$
Part III	$2\cdot23\leqslant\gamma_w\leqslant8\cdot54$

VIII. SUMMARY

The problem of the reliability of mental tests has been discussed using the theory of testing statistical hypotheses, and of estimation, devised and developed by Neyman and Pearson. Four separate problems have been considered:

(1) to determine if there is a significant trial effect,

(2) to determine if the mental test actually measures the capacity of the individuals tested,

(3) to estimate the trial effect if it exists, and

(4) to obtain an estimate of the relative importance of the random errors of measurement, with respect to the true measurement of the capacity of the individual, in determining the score of the individual on the mental test.

A new concept, the sensitivity of a mental test, has been introduced as a measure of reliability in the solution of problems (2) and (4). The results for three mental tests, Comprehensive Examinations in Biological Sciences, have been analysed and discussed.

In conclusion I wish to thank Dr J. Neyman for the help which he has given me throughout the course of the work.

REFERENCES

(1) NEYMAN, J. (1935). "Sur la vérification des hypothèses statistiques composées." *Bull. Soc. math. Fr.* LXIII.

(2) —— (1937). "Outline of a theory of estimation based on the classical theory of probability." *Philos. Trans.* (in the Press).

(3) NEYMAN, J. & PEARSON, E. S. (1936). "Sufficient statistics and uniformly most powerful tests of statistical hypotheses." *Statist. Res. Mem.* Department of Applied Statistics, University College, London, I.

(4) PEARSON, K. (1934). *Tables of the Incomplete Beta-Function.*

(5) FISHER, R. A. (1932). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

(6) GOODENOUGH, FLORENCE L. (1936). "A critical note on the use of the term 'reliability' in mental measurement." *J. Educ. Psychol.* XXVII, 173–8.

(7) JOHNSON, P. O. & NEYMAN, J. (1936). "Tests of certain linear hypotheses and their application to some educational problems." *Statist. Res. Mem.* Department of Applied Statistics, University College, London, I.

(8) THOULESS, R. H. (1935). "Test unreliability and function fluctuation." *Brit. J. Psychol.* XXVI, 325–43.

(9) JACKSON, R. W. B. (1939). "Some problems of testing statistical hypotheses and estimation relating to the question of the relative accuracy of measurements." *Statist. Res. Mem.*, III.

(*Manuscript received* 25 *August* 1937)

THE INFLUENCE OF MULTIVARIATE SELECTION ON THE FACTORIAL ANALYSIS OF ABILITY

BY GODFREY H. THOMSON (SECTIONS I–V)
AND WALTER LEDERMANN (SECTIONS VI–VIII)
Moray House, University of Edinburgh

I. *Introduction* (pp. 288–289).
II. *Aitken's formula for multivariate selection* (pp. 289–290).
III. *(Minimum) rank unchanged in the 'indirect' quadrants* (pp. 290–291).
IV. *Creation and annihilation of group factors* (pp. 291–292).
V. *The influence of selection on the factors themselves* (pp. 292–297).
VI. *An algebraical proof of the theorem on multivariate selection* (pp. 297–304).
  (*a*) *The induced correlations between factors* (pp. 297–299).
  (*b*) *Analysis of the tests after selection* (pp. 299–304).
VII. *The number of induced group factors* (pp. 304–305).
VIII. *Application to the theory of univariate selection* (pp. 305–306).
References (p. 306).

I. INTRODUCTION

IN previous articles (1, 2) the influence of univariate selection on factorial analysis was considered, and it was shown that such selection leaves the rank of the 'reduced'* matrix of correlations unchanged although the loadings of all the factors are altered, i.e. one and the same test has different saturations with *g*, etc., in the population and the sample.

In view of what is to follow it is desirable to re-emphasize here what is stressed in the concluding sections of the previous paper, namely, that there is no real difference between population and sample in these considerations, and the relationships found work both ways. The word 'selection' suggests that all the individuals in the sample are contained among the individuals in the population. Even if that were so, a sample could be selected which had a larger standard deviation than the population. And in evolution, where in a certain sense one generation is a sample of the preceding, it may nevertheless outnumber its parents. It is better therefore to think of both sample and population as being samples from a hypothetical superpopulation of 'might-have-beens' (see (3), p. 183, quoted in (4), p. 188). It is clear then that there are as many analyses of a test as there are sample populations we can use it on, and that, mathematically at least, no one of these analyses has any claim to be considered the 'true' one. Univariate selection, however, does not alter the *number* of factors.

---

* By the reduced correlation matrix is meant in Thurstone's terminology the matrix with those quantities inserted in the principal diagonal which reduce its rank to a minimum. These quantities are called 'communalities', their defects from unity give the squares of the loadings of the specific factors, and the analysis for common factors is carried out on the reduced matrix. Throughout this paper I shall use the term 'minimum attainable rank' or for brevity 'minimum rank' to mean the lowest rank to which a matrix can be reduced by changing its leading diagonal.

If now we turn to multivariate selection, where several variates are explicitly selected, we find that of course 'new factors' can be created, or old ones destroyed; but we shall see that these changes in the number of factors are only among the 'selected' variates, and that the remaining variates can still be analysed into the old number of factors, although with new loadings, and not in general identical in nature with the old factors.

II. AITKEN'S FORMULA FOR MULTIVARIATE SELECTION

Pearson's theory of multivariate selection has received considerable clarification from the formula into which it has been condensed by two recent papers of A. C. Aitken (5, 6). If $n=p+q$ tests are normalized their intercorrelations may be expressed in the $n \times n$ matrix

$$\begin{bmatrix} R_{pp} & R_{pq} \\ R_{qp} & R_{qq} \end{bmatrix}. \qquad \ldots\ldots(1)$$

If now the variances and 'covariances'* of the first p tests are directly modified by 'selection' (in the wide sense) to new values V_{pp} then the rest of the matrix will be sympathetically modified to the form

$$\begin{bmatrix} V_{pp} & V_{pp}R_{pp}^{-1}R_{pq} \\ R_{qp}R_{pp}^{-1}V_{pp} & R_{qq}-R_{qp}(R_{pp}^{-1}-R_{pp}^{-1}V_{pp}R_{pp}^{-1})\,R_{pq} \end{bmatrix}. \qquad \ldots\ldots(2)$$

This is Aitken's formula: it is certainly applicable when the distributions are all normal, and it is believed to be applicable over a considerably wider range, though I understand that there is some doubt as to how wide this range is.

---

* The word has presumably got too good a start to be recalled. But it already had a defined meaning in a closely allied branch of mathematics and its new use may lead to confusion. The new variances are referred to the new means if these have changed.

III. (MINIMUM) RANK UNCHANGED IN THE 'INDIRECT' QUADRANTS

As a preliminary example let us take the same matrix of correlations between five hierarchical tests as was used in (1), with an additional test added, namely,

$$\left[\begin{array}{cc|cccc} 1{\cdot}00 & 0{\cdot}72 & 0{\cdot}63 & 0{\cdot}54 & 0{\cdot}45 & 0{\cdot}36 \\ 0{\cdot}72 & 1{\cdot}00 & 0{\cdot}56 & 0{\cdot}48 & 0{\cdot}40 & 0{\cdot}32 \\ \hline 0{\cdot}63 & 0{\cdot}56 & 1{\cdot}00 & 0{\cdot}42 & 0{\cdot}35 & 0{\cdot}28 \\ 0{\cdot}54 & 0{\cdot}48 & 0{\cdot}42 & 1{\cdot}00 & 0{\cdot}30 & 0{\cdot}24 \\ 0{\cdot}45 & 0{\cdot}40 & 0{\cdot}35 & 0{\cdot}30 & 1{\cdot}00 & 0{\cdot}20 \\ 0{\cdot}36 & 0{\cdot}32 & 0{\cdot}28 & 0{\cdot}24 & 0{\cdot}20 & 1{\cdot}00 \end{array}\right], \qquad \text{......(3)}$$

and let us modify by intentional selection not only as there the first test, but the first two *and their covariance*, by changing the north-west quadrant of the matrix to

$$\begin{bmatrix} 0{\cdot}36 & 0{\cdot}30 \\ 0{\cdot}30 & 0{\cdot}36 \end{bmatrix}, \qquad \text{......(4)}$$

that is to say, we have selected a sample so that the individuals forming it are much less scattered both in test 1 and test 2 (variances 0·36 and 0·36 instead of 1·00) and also show a higher correlation between their scores in these two tests (0·83 instead of 0·72). Aitken's formula enables us then to write down what the variances and covariances of the remaining three tests become by reason of the correlations with the first two, namely,

$$\left[\begin{array}{cc|cccc} 0{\cdot}360 & 0{\cdot}300 & 0{\cdot}236 & 0{\cdot}202 & 0{\cdot}169 & 0{\cdot}135 \\ 0{\cdot}300 & 0{\cdot}360 & 0{\cdot}220 & 0{\cdot}189 & 0{\cdot}158 & 0{\cdot}126 \\ \hline 0{\cdot}236 & 0{\cdot}220 & 0{\cdot}740 & 0{\cdot}196 & 0{\cdot}163 & 0{\cdot}131 \\ 0{\cdot}202 & 0{\cdot}189 & 0{\cdot}196 & 0{\cdot}808 & 0{\cdot}140 & 0{\cdot}113 \\ 0{\cdot}169 & 0{\cdot}158 & 0{\cdot}163 & 0{\cdot}140 & 0{\cdot}868 & 0{\cdot}094 \\ 0{\cdot}135 & 0{\cdot}126 & 0{\cdot}131 & 0{\cdot}113 & 0{\cdot}094 & 0{\cdot}915 \end{array}\right], \qquad \text{......(5)}$$

which on being normalized gives the matrix of correlations

$$\left[\begin{array}{cc|cccc}
1{\cdot}000 & 0{\cdot}833 & 0{\cdot}457 & 0{\cdot}375 & 0{\cdot}302 & 0{\cdot}235 \\
0{\cdot}833 & 1{\cdot}000 & 0{\cdot}426 & 0{\cdot}351 & 0{\cdot}283 & 0{\cdot}219 \\
\hline
0{\cdot}457 & 0{\cdot}426 & 1{\cdot}000 & 0{\cdot}253 & 0{\cdot}203 & 0{\cdot}159 \\
0{\cdot}375 & 0{\cdot}351 & 0{\cdot}253 & 1{\cdot}000 & 0{\cdot}167 & 0{\cdot}131 \\
0{\cdot}302 & 0{\cdot}283 & 0{\cdot}203 & 0{\cdot}167 & 1{\cdot}000 & 0{\cdot}105 \\
0{\cdot}235 & 0{\cdot}219 & 0{\cdot}159 & 0{\cdot}131 & 0{\cdot}105 & 1{\cdot}000
\end{array}\right]. \qquad \ldots\ldots(6)$$

It is this last matrix which would be found by an experimenter who *began* with the sample. And examination shows that *it is still hierarchical in the three indirectly selected quadrants*. The only correlation which breaks the hierarchy is the 0·833 which we made intentionally of that size by selection. Similarly the matrix of variances and covariances is hierarchical except for the 0·300 of tests 1 and 2.

This discovery I was rapidly able to confirm by other examples of larger order and of higher rank. In every case, the minimum rank of the matrix, that is to say the rank to which it can be reduced by placing suitable 'communalities' in the leading diagonal, remained the same in the three quarters of the matrix *indirectly* selected, and was only changed in the directly selected quadrant.

IV. CREATION AND ANNIHILATION OF GROUP FACTORS

In our simple example we note that since the directly selected quadrant no longer conforms to the hierarchy, *a new group factor has been created by the selection*, and would be presumed to exist by an experimenter who began with what we are calling the 'sample'. The original correlation matrix could be analysed as follows:

$$\left.\begin{aligned}
z_1 &= 0{\cdot}9f + 0{\cdot}436s_1, \\
z_2 &= 0{\cdot}8f + 0{\cdot}600s_2, \\
z_3 &= 0{\cdot}7f + 0{\cdot}714s_3, \\
z_4 &= 0{\cdot}6f + 0{\cdot}800s_4, \\
z_5 &= 0{\cdot}5f + 0{\cdot}866s_5, \\
z_6 &= 0{\cdot}4f + 0{\cdot}917s_6,
\end{aligned}\right\} \qquad \ldots\ldots(7)$$

the new correlation matrix as follows:

$$\left.\begin{aligned} z_1 &= 0{\cdot}824f_1 + 0{\cdot}447f_1 + 0{\cdot}348s_1, \\ z_2 &= 0{\cdot}769f_1 + 0{\cdot}447f_2 + 0{\cdot}457s_2, \\ z_3 &= 0{\cdot}555f_1 \qquad\qquad\quad + 0{\cdot}832s_3, \\ z_4 &= 0{\cdot}456f_1 \qquad\qquad\quad + 0{\cdot}890s_4, \\ z_5 &= 0{\cdot}366f_1 \qquad\qquad\quad + 0{\cdot}931s_5, \\ z_6 &= 0{\cdot}286f_1 \qquad\qquad\quad + 0{\cdot}958s_6. \end{aligned}\right\} \qquad \ldots\ldots(8)$$

If the 'sample' is the whole of the present generation then the 'new' group factor f_2 is presumably as real as any other group factor. The population from which the sample is 'selected' may be inaccessible to us. It may be the preceding generation from whom our sample is descended. It may be the present generation extended by all those who died in infancy, or even by all those who were never born. It becomes abundantly clear on taking thought that *all* populations are samples.

The process of selection, and Aitken's formula, is reflexive, and our example can be worked backwards, illustrating the disappearance of a 'factor' through evolution or through actual human selection. The 'purification' of a hierarchical battery of tests has usually been carried out by evicting tests from the test sample. It could also be carried out by evicting persons from the population of subjects, a process which could be engineered so as to make the erring parts of the matrix conform to the hierarchy, without injuring the hierarchical perfection of the remainder.

V. THE INFLUENCE OF SELECTION ON THE FACTORS THEMSELVES

Shortly afterwards it occurred to me that a still wider use of Aitken's formula can be made, which throws yet more light on the nature of factors. The latter themselves are variates exactly as the tests are, and the analysis of the tests has given us the correlations of the factors with the tests and with each other (we are considering orthogonal factors but oblique factors could also be handled in the same way). If we write the factor correlations (both of group and specific factors) down exactly as though the factors were a continuation of the series of tests we obtain the matrix

$$\left[\begin{array}{c|cc} R_{pp} & R_{pq} & R_{pf} \\ \hline R_{qp} & R_{qq} & R_{qf} \\ R_{fp} & R_{fq} & R_{ff} \end{array}\right], \qquad \ldots\ldots(9)$$

and the result of changing R_{pp} by selection to V_{pp} will be to give the matrix

$$\begin{bmatrix} V_{pp} & V_{pq} & V_{pf} \\ V_{qp} & V_{qq} & V_{qf} \\ V_{fp} & V_{fq} & V_{ff} \end{bmatrix}, \qquad \text{......(10)}$$

where

$$\left[\begin{array}{c|c} V_{pp} & V_{pq} \\ \hline V_{qp} & V_{qq} \end{array}\right]$$

is the same as the matrix (2), and

$$\left.\begin{aligned} V_{pf} &= V_{pp} R_{pp}^{-1} R_{pf}, \\ V_{qf} &= R_{qf} - R_{qp} R_{pp}^{-1} R_{pf} + R_{qp} R_{pp}^{-1} V_{pp} R_{pp}^{-1} R_{pf}, \\ V_{ff} &= R_{ff} - R_{fp} R_{pp}^{-1} R_{pf} + R_{fp} R_{pp}^{-1} V_{pp} R_{pp}^{-1} R_{pf}, \end{aligned}\right\} \qquad \text{......(11)}$$

and if the factors are originally orthogonal $R_{ff} = I$. Note that the actual calculations need not be divided up but can be done all in a piece.

The letter f above refers to *all* the factors. We can, however, if we wish divide the matrix up still further. If we henceforward restrict f to mean common or group factors, s to mean the specifics of the directly selected variates, and t to mean the specifics of the other variates, we can write the original matrix as

$$\left[\begin{array}{c|cccc} R_{pp} & R_{pq} & R_{pf} & R_{ps} & R_{pt} \\ \hline R_{qp} & R_{qq} & R_{qf} & R_{qs} & R_{qt} \\ R_{fp} & R_{fq} & R_{ff} & R_{fs} & R_{ft} \\ R_{sp} & R_{sq} & R_{sf} & R_{ss} & R_{st} \\ R_{tp} & R_{tq} & R_{tf} & R_{ts} & R_{tt} \end{array}\right], \qquad \text{......(12)}$$

(where in the case of orthogonal factors R_{ff}, R_{ss}, and R_{tt} are all unit matrices, and R_{fs}, R_{ft}, R_{st} and their transposes zero matrices). The matrices after selection in the p variates will be represented by the same pattern with V instead of R (the suffixes unchanged), and it is convenient to mention here that Dr Walter Ledermann (who will use this arrangement in the later sections of this paper) has shown that Aitken's formula can be compressed into the statement, true for any number of cells in the above kind of dissection, that each submatrix (including V_{pp} itself) is given by the formula

$$\left.\begin{aligned} V_{\alpha\beta} &= R_{\alpha\beta} - R_{\alpha p} E_{pp} R_{p\beta} \quad (\alpha,\ \beta = p,\ q,\ f, \ldots), \\ \text{where} \qquad E_{pp} &= R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1}. \end{aligned}\right\} \qquad \text{......(12}a\text{)}$$

A numerical example will best illustrate the conclusions that can be drawn from this wider application of the selection formula. Let the original matrix of correlations of five tests be as shown in (13), together with the correlations of these tests with their two common factors and five specifics, all the factors being orthogonal. Table (13) is in fact a complete representation of what Thurstone calls a 'structure'.

	p		*q*			*f*		*s*		*t*		
p	1·00	0·46	0·59	0·36	0·41	0·70	0·40	0·59	.	.	.	.
	0·45	1·00	0·36	0·26	0·23	0·60	0·10	.	0·79	.	.	.
q	0·59	0·36	1·00	0·32	0·45	0·50	0·60	.	.	0·62	.	.
	0·36	0·26	0·32	1·00	0·22	0·40	0·20	.	.	.	0·89	.
	0·41	0·23	0·45	0·22	1·00	0·30	0·50	.	.	.	.	0·81
f	0·70	0·60	0·50	0·40	0·30	1·00	.	.	.	.	.	.
	0·40	0·10	0·60	0·20	0·50	.	1·00	.	.	.	.	.
s	0·59	.	.	.	.	.	.	1·00	.	.	.	.
	.	0·79	.	.	.	.	.	.	1·00	.	.	.
t	.	.	0·62	.	.	.	.	.	.	1·00	.	.
	.	.	.	0·89	.	.	.	.	.	.	1·00	.
	.	.	.	.	0·81	.	.	.	.	.	.	1·00

.....(13)

The numbers in this table are supposed to be exact, except that the loadings of the specifics should be exactly the square roots of 0·35, 0·63, 0·39, 0·80, 0·66. In future tables only two places of decimals are printed, for clearness and economy of space, but the calculations were of course carried further.

Now let the variances and covariances of the first two tests be changed to the values shown in (14), upon which the rest of the matrix, *including the correlations of the factors*, will change sympathetically to the values there shown. The matrix of new correlations can at once be written down from (14), but we shall omit it to save space.

The first point of interest which arises from an inspection of this table (14) is that *the originally orthogonal factors are not now all orthogonal.* Those of them which are concerned in the composition of the directly selected variates have become correlated.

The second point of interest is that the minimum attainable rank of the matrix remains unchanged within the square of the tests and their common factors, except for the corner square V_{pp}. Indeed the area of rank r (rank 2 in our example) has extended to include the cross-variances of the specifics of the selected tests with the other tests and the common factors, an area which formerly was blank, so that *the area of minimum attainable rank* r *forms a broad cross on the table,* as marked.

	p		*q*			*f*		*s*		*t*		
p	0·36	0·36	0·23	0·15	0·16	0·33	0·12	0·15	0·20	.	.	.
	0·36	0·64	0·27	0·19	0·17	0·43	0·09	0·05	0·48	.	.	.
q	0·23	0·27	0·80	0·20	0·31	0·28	0·45	−0·24	0·07	0·62	.	.
	0·15	0·19	0·20	0·93	0·14	0·26	0·12	−0·13	0·02	.	0·89	.
	0·16	0·17	0·31	0·14	0·90	0·15	0·39	−0·17	0·06	.	.	0·81
f	0·33	0·43	0·28	0·26	0·15	0·74	−0·15	−0·22	−0·01	.	.	.
	0·12	0·09	0·45	0·12	0·39	−0·15	0·88	−0·21	0·12	.	.	.
s	0·15	0·05	−0·24	−0·13	−0·17	−0·22	−0·21	0·65	0·26	.	.	.
	0·20	0·48	0·07	0·02	0·06	−0·01	0·12	0·26	0·59	.	.	.
t	.	.	0·62	.	.	.	.	.	.	1·00	.	.
	.	.	.	0·89	.	.	.	.	.	.	1·00	.
	.	.	.	.	0·81	.	.	.	.	.	.	1·00

......(14)

Since the test square itself is still of minimum attainable rank r except for V_{pp}, it can be re-analysed into the same number of *orthogonal* common factors as before, plus new group factors d confined to the p directly selected tests, as is shown in (15) giving variances and covariances, and in (16) giving correlations.

There are seen in (15) and (16) two new orthogonal group factors e in place of the old and now oblique factors f; a new group factor d running through the directly selected tests only; and new orthogonal specifics y to replace the specifics s which have become oblique. This is the manner in which an experimenter who began with the sample would analyse the

$$\begin{array}{c|cc|ccc|cc|c|cc|ccc}
 & p & & q & & & e & & d & y & & t & & \\
\hline
p & 0{\cdot}36 & 0{\cdot}36 & 0{\cdot}23 & 0{\cdot}15 & 0{\cdot}16 & 0{\cdot}37 & 0{\cdot}21 & 0{\cdot}37 & 0{\cdot}20 & . & . & . & . \\
 & 0{\cdot}36 & 0{\cdot}64 & 0{\cdot}27 & 0{\cdot}19 & 0{\cdot}17 & 0{\cdot}49 & 0{\cdot}20 & 0{\cdot}37 & . & 0{\cdot}48 & . & . & . \\
\hline
q & 0{\cdot}23 & 0{\cdot}27 & 0{\cdot}80 & 0{\cdot}20 & 0{\cdot}31 & 0{\cdot}33 & 0{\cdot}55 & . & . & . & 0{\cdot}62 & . & . \\
 & 0{\cdot}15 & 0{\cdot}19 & 0{\cdot}20 & 0{\cdot}93 & 0{\cdot}14 & 0{\cdot}31 & 0{\cdot}18 & . & . & . & . & 0{\cdot}89 & . \\
 & 0{\cdot}16 & 0{\cdot}17 & 0{\cdot}31 & 0{\cdot}14 & 0{\cdot}90 & 0{\cdot}17 & 0{\cdot}46 & . & . & . & . & . & 0{\cdot}81 \\
\hline
e & 0{\cdot}37 & 0{\cdot}49 & 0{\cdot}33 & 0{\cdot}31 & 0{\cdot}17 & 1{\cdot}00 & . & . & . & . & . & . & . \\
 & 0{\cdot}21 & 0{\cdot}20 & 0{\cdot}55 & 0{\cdot}18 & 0{\cdot}46 & . & 1{\cdot}00 & . & . & . & . & . & . \\
\hline
d & 0{\cdot}37 & 0{\cdot}37 & . & . & . & . & . & 1{\cdot}00 & . & . & . & . & . \\
\hline
y & 0{\cdot}20 & . & . & . & . & . & . & . & 1{\cdot}00 & . & . & . & . \\
 & . & 0{\cdot}48 & . & . & . & . & . & . & . & 1{\cdot}00 & . & . & . \\
\hline
t & . & . & 0{\cdot}62 & . & . & . & . & . & . & . & 1{\cdot}00 & . & . \\
 & . & . & . & 0{\cdot}89 & . & . & . & . & . & . & . & 1{\cdot}00 & . \\
 & . & . & . & . & 0{\cdot}81 & . & . & . & . & . & . & . & 1{\cdot}00
\end{array} \qquad \ldots\ldots(15)$$

tests (though he can 'rotate' his loadings) if he were wedded to orthogonal factors.

The concept of orthogonal factors is indeed even more fluid and dependent upon the sample than I realized in my former paper on univariate selection. After univariate selection the matrix can, it is true, be analysed into an orthogonal pattern of exactly the same form (position of zeros) as before, though the loadings have changed. But some at least of the factors have also changed, namely those concerned in the selected variate, for the previous factors are now correlated and the new ones are orthogonal. In the special case of *univariate* selection on a *hierarchical* matrix the change in nature can I think be confined to the specific of the directly selected test.

The discoveries described in the above were made partly empirically, partly by a loose use of various algebraic and geometric tools and a good deal of intuition. I therefore invited Dr Walter Ledermann to submit

	p		q			e		d	y		t		
p	1·00	0·75	0·44	0·26	0·28	0·62	0·34	0·62	0·34	.	.	.	.
	0·75	1·00	0·37	0·24	0·23	0·61	0·24	0·46	.	0·60	.	.	.
q	0·44	0·37	1·00	0·23	0·36	0·36	0·62	.	.	.	0·70	.	.
	0·26	0·24	0·23	1·00	0·15	0·32	0·19	.	.	.	.	0·93	.
	0·28	0·23	0·36	0·15	1·00	0·18	0·48	.	.	.	.	.	0·86
e	0·62	0·61	0·36	0·32	0·18	1·00	.	.	.	.	.	.	.
	0·34	0·24	0·62	0·19	0·48	.	1·00	.	.	.	.	.	.
d	0·62	0·46	.	.	.	.	.	1·00	.	.	.	.	.
y	0·34	.	.	.	.	.	.	.	1·00	.	.	.	.
	.	0·60	.	.	.	.	.	.	.	1·00	.	.	.
t	.	.	0·70	.	.	.	.	.	.	.	1·00	.	.
	.	.	.	0·93	.	.	.	.	.	.	.	1·00	.
	.	.	.	.	0·86	.	.	.	.	.	.	.	1·00

......(16)

them to a rigorous examination and give exact proofs, and this, with some additional matter, he gives in the remaining sections of this paper, which are from his pen.

VI. AN ALGEBRAICAL PROOF OF THE THEOREM ON MULTIVARIATE SELECTION

(a) *The induced correlations between factors*

This section contains a formal proof of Prof. Godfrey H. Thomson's theorem on multivariate selection which has been set forth and illustrated by him in the first part of this paper. The theorem can be stated as follows (cf. p. 289):

If a team of $n=p+q$ *tests has been resolved into* r *general and* n *specific factors which are normalized and mutually orthogonal, and if the variances and covariances of* p *of these tests are changed by selection, the mean (if also* *changed) being again taken as the origin, then the battery can still be analysed into* r *generals and* n *specifics, but, in addition, a certain number* l *of group factors will appear which, however, run through the* p *selected tests only. Further these* r + n + l *factors will again be uncorrelated and normalized.*

Let $z = (z_p, z_q)$ be the vector of the $n = p + q$ tests, where z_p denotes the directly selected tests and where z_q refers to the remaining tests. We assume that the battery of tests has been analysed thus

$$\left.\begin{aligned} z_p &= R_{pf} f + R_{ps} s \quad, \\ z_q &= R_{qf} f \quad + R_{qt} t, \end{aligned}\right\} \qquad \text{......(17)}$$

where f is a vector of r common factors, and s and t represent p and q specifics respectively, corresponding to the two groups of tests. The elements of the matrices R_{pf}, R_{ps}, R_{qf} and R_{qt} are the loadings of the general and specific factors and, in particular, R_{ps} and R_{qt} are diagonal matrices. The analysis (17) refers to a hypothetical superpopulation (see p. 289) which contains both the parent and the selected population as samples. Equations (17) therefore hold for the scores of each person under consideration, to whichever sample he may belong. We have suppressed a second subscript which would specify the individual person, and we shall rather regard each element of the column vectors $z_p, z_q, f, \ldots$ as a row vector with N components corresponding to the number of persons in the sample. Moreover, we shall assume that the scale is chosen in such a way that the correlations are simply given by the scalar products of these row vectors. Thus we have

$$\left.\begin{aligned} R_{pf} &= z_p f', \quad R_{ps} = z_p s', \quad R_{pt} = z_p t' = 0, \\ R_{qf} &= z_q f', \quad R_{qs} = z_q s' = 0, \quad R_{qt} = z_p t', \end{aligned}\right\} \qquad \text{......(18)}$$

since z_p is independent of t, and z_q is independent of s, and remembering that the factors themselves are uncorrelated and normalized,

$$R_{ff} = ff' = I, \quad R_{ss} = ss' = I, \quad R_{tt} = tt' = I,$$

$$R_{fs} = fs' = 0, \quad R_{ft} = ft' = 0, \quad R_{st} = st' = 0.$$

After selection, the variances and covariances are

$$V_{\alpha\beta} = R_{\alpha\beta} - R_{\alpha p}(R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1})\, R_{p\beta} = R_{\alpha\beta} - R_{\alpha p} E_{pp} R_{p\beta}, \qquad \text{......(12a)}$$

where α and β take the values p, q, f, s, t (see p. 298). In particular, by means of (18), we find that

$$\left.\begin{aligned}&V_{st}=0, \quad V_{ft}=0, \quad V_{tt}=I,\\ &V_{fs}=-R_{fp}(R_{pp}^{-1}-R_{pp}^{-1}V_{pp}R_{pp}^{-1})\,R_{ps}=-R_{fp}E_{pp}R_{ps},\\ &V_{ff}=I-R_{fp}(R_{pp}^{-1}-R_{pp}^{-1}V_{pp}R_{pp}^{-1})\,R_{pf}=I-R_{fp}E_{pp}R_{pf},\\ &V_{pf}=V_{pp}R_{pp}^{-1}R_{pf}.\end{aligned}\right\} \quad \ldots\ldots(19)$$

(b) *Analysis of the tests after selection*

We shall use a circumflex accent to denote the scores (referred to the new means) of the persons in the selected population. From the general relations (17) it immediately follows that

$$\left.\begin{aligned}\hat{z}_p&=R_{pf}\hat{f}+R_{ps}\hat{s} &&,\\ \hat{z}_q&=R_{qf}\hat{f} &&+R_{qt}\hat{t},\end{aligned}\right\} \quad \ldots\ldots(20)$$

where the coefficients $R_{pf}, \ldots,$ etc., are unaltered. We shall, however, assume that the new variances and covariances are again equal to the scalar products of the score vectors. This implies only a change of units; in fact, we have to multiply (17) throughout by $\sqrt{N/N'}$, where N' is the number of persons in the selected population. Thus we have

$$\hat{z}_p\hat{z}_p'=V_{pp}, \quad \hat{z}_p\hat{f}'=V_{pf}, \ldots, \text{ etc.} \quad \ldots\ldots(21)$$

In the selected population, the factors $\hat{f}$, $\hat{s}$, $\hat{t}$ are no longer orthogonal, and our aim is now to replace them by new factors, which shall be uncorrelated and normalized. It will be seen that this can, in general, be achieved only if further dimensions are added to the common-factor-space.

First, we shall split up $\hat{s}$ into a part which lies in the original common-factor-space and a part perpendicular to that space, thus

$$\hat{s}=X\hat{f}+\hat{u}, \quad \ldots\ldots(22)$$

where

$$\hat{u}\hat{f}'=0. \quad \ldots\ldots(23)$$

On multiplying (22) by $\hat{f}'$ we find

$$\hat{s}\hat{f}'=X\hat{f}\hat{f}',$$

or, by (21),

$$V_{sf}=XV_{ff},$$

whence

$$X=V_{sf}V_{ff}^{-1}.$$

Substituting this in (22) we get

$$\hat{s}=V_{sf}V_{ff}^{-1}\hat{f}+\hat{u}. \quad \ldots\ldots(24)$$

In general, we shall have

$$V_{ff} \neq I \quad \text{(see (19))},$$

i.e. the variates f will be correlated in the selected population. But the matrix V_{ff} is a 'Gramian' matrix (being a correlation matrix) and is therefore capable of being factorized, thus

$$V_{ff} = KK', \qquad \text{......(25)}$$

where K is a square matrix with real elements. There are various methods by which the matrix K can be computed, e.g. L. L. Thurstone's 'diagonal method' ((7), p. 78), but the solution is not unique, a right-hand orthogonal factor being arbitrary. This indeterminacy is, however, inherent in every factor problem. After deciding on a particular solution K of (25) we put

$$\hat{f} = K\hat{e}, \qquad \text{......(26)}$$

where $\hat{e}$ represents r new group factors which are now uncorrelated and normalized; for we have

$$\hat{e}\hat{e}' = K^{-1}\hat{f}\hat{f}'K^{-1\prime} = K^{-1}V_{ff}K^{-1\prime},$$

whence, by (25),

$$\hat{e}\hat{e}' = I. \qquad \text{......(27)}$$

Substituting (24) and (26) in (20) we get

$$\left.\begin{aligned} \hat{z}_p &= (R_{pf} + R_{ps}V_{sf}V_{ff}^{-1})\,K\hat{e} + R_{ps}\hat{u}, \\ \hat{z}_q &= \qquad R_{qf}K\hat{e} \qquad\qquad + R_{qt}\hat{t}. \end{aligned}\right\} \qquad \text{......(28)}$$

The coefficient of $\hat{e}$ in the first equation can be more conveniently written as follows

$$\left.\begin{aligned} L &= (R_{pf} + R_{ps}V_{sf}V_{ff}^{-1})\,K = R_{pf}K + R_{ps}V_{sf}(KK')^{-1}K \\ &\qquad = R_{pf}K + R_{ps}V_{sf}K^{-1\prime} \quad \text{(by (25))}, \\ L &= (R_{pf}KK' + R_{ps}V_{sf})\,K^{-1\prime} \\ &= [R_{pf}(I - R_{fp}E_{pp}R_{pf}) - R_{ps}R_{sp}E_{pp}R_{pf}]K^{-1\prime} \quad \text{(by (19))}. \end{aligned}\right\} \text{.......(29)}$$

Now, the correlation matrices $R_{\alpha\beta}$ are not independent of one another, in virtue of the linear equations (17). Multiplying the first of these equations by its transpose we obtain

$$z_p z_p' = (R_{pf}f + R_{ps}s)\,(f'R_{fp} + s'R_{sp}),$$

and since the factors are mutually orthogonal and standardized in the original population

$$R_{pp} = R_{pf}R_{fp} + R_{ps}^{2}, \qquad \text{......(30)}$$

remembering that $R_{ps} = R_{sp}$ is a diagonal matrix. Using (30) to eliminate the product $R_{pf}R_{fp}$ we can write (29) in the form

$$L = [R_{pf} - (R_{pp} - R_{ps}^2)\, E_{pp}R_{pf} - R_{ps}^2 E_{pp} R_{pf}]\, K^{-1\prime},$$

and on cancelling the last two terms

$$L = [R_{pf} - R_{pp}E_{pp}R_{pf}]\, K^{-1\prime}.$$

Finally, on substituting for E_{pp} from (12*a*), p. 298, we get

$$L = [R_{pf} - R_{pp}(R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1})\, R_{pf}]\, K^{-1\prime},$$

which readily simplifies to

$$L = V_{pp}R_{pp}^{-1}R_{pf}K^{-1\prime}.$$

Substituting this in (28) and putting, for brevity,

$$\hat{w} = R_{ps}\hat{u}, \qquad \text{......(31)}$$

we obtain

$$\left.\begin{aligned} \hat{z}_p &= V_{pp}R_{pp}^{-1}R_{pf}K^{-1\prime}\hat{e} + \hat{w}, \\ \hat{z}_q &= \qquad R_{qf}K\hat{e} \qquad + R_{qt}\hat{t}. \end{aligned}\right\} \qquad \text{......(32)}$$

The test scores are now expressed in terms of three sets of factors, viz. $\hat{e}$, $\hat{t}$ and $\hat{w}$. Factors belonging to different sets are orthogonal, for we have

$$\hat{e}\hat{t}' = K^{-1}\hat{f}\hat{t}' = K^{-1}V_{ft} = 0, \quad \text{by (26) and (19),}$$

$$\hat{e}\hat{w}' = K^{-1}\hat{f}\hat{u}'R_{sp} = K^{-1}(\hat{u}\hat{f}')'R_{sp} = 0, \quad \text{by (23),}$$

$$\hat{t}\hat{w}' = \hat{t}\hat{u}'R_{sp} = \hat{t}(\hat{s} - V_{sf}V_{ff}^{-1}\hat{f})'R_{sp} = \hat{t}\hat{s}' - \hat{t}\hat{f}'V_{ff}^{-1}V_{fs}R_{sp} = 0,$$

since, by (19),

$$V_{st} = V_{ft} = 0.$$

Moreover, the various components of $\hat{e}$ and of $\hat{t}$ are within themselves uncorrelated and normalized, since by (27) and (19)

$$\hat{e}\hat{e}' = V_{ee} = I,$$

$$\hat{t}\hat{t}' = V_{tt} = I.$$

We have, therefore, reached our goal except for the p factors $\hat{w}$ which will, in general, not be orthogonal. However, we could orthogonalize them by a process similar to that which was used to transform $\hat{f}$ into $\hat{e}$, and we should then obtain p new orthogonal group factors. But in accordance with the principle of minimizing the number of group factors at the expense of specific factors we shall endeavour to resolve $\hat{w}$ into as small a number l of group factors as possible and into p specifics (see § VII below), thus

$$\hat{w} = A_{wd}\hat{d} + B_{wy}\hat{y}, \qquad \text{......(33)}$$

where $\hat{d}$ represents the l group factors and $\hat{y}$ the p specifics; the elements of the matrices A_{pd} and B_{py} are the corresponding loadings, the matrix B_{py} being a non-singular diagonal matrix.

Now, (33) constitutes Thurstone's factor problem for the p fictitious tests $\hat{w}$ and the required analysis can be carried out by the usual methods. The matrix of variances and covariances of $\hat{w}$ which determines the factorial analysis of $\hat{w}$ will be denoted by V_{ww}, and an explicit expression for it will be derived forthwith. It should be noted that $\hat{w}$ is not in standard measure, but the process which turns V_{ww} into the actual correlation matrix (i.e. dividing each row and column by the respective standard deviation) only slightly modifies Thurstone's method of factorizing a team of tests. The number of new group factors $\hat{d}$ is determined by the 'minimum rank' of V_{ww} and after substituting from (33) in (32) we see that the number of dimensions of the total factor space has now risen to

$$r+p+q+l=r+n+l,$$

i.e. l dimensions more than before selection. We can therefore place the additional l group factors at right angles to each of the old factors $\hat{f}$, $\hat{s}$, $\hat{t}$, or $\hat{e}$, $\hat{w}$, $\hat{t}$, since the latter are linear functions of the former factors (see (26), (31) and (24)). From (33) it then follows that the new specifics $\hat{y}$ are likewise orthogonal to $\hat{e}$ and $\hat{t}$. For after post-multiplying (33) by $\hat{e}'$ we get

$$\hat{w}\hat{e}'=A_{wd}\hat{d}\hat{e}'+B_{wv}\hat{y}\hat{e}',$$

whence

$$\hat{y}\hat{e}'=0,$$

since $\hat{w}\hat{e}'=0$ and $\hat{d}\hat{e}'=0$ and since B_{wd} is non-singular. In the same way it can be shown that

$$\hat{y}\hat{t}'=0.$$

Thus from (32) and (33) we get

$$\left.\begin{aligned}\hat{z}_p&=V_{pp}R_{pp}^{-1}R_{pf}K^{-1\prime}\hat{e}+A_{wd}\hat{d}+B_{wd}\hat{y},\\ \hat{z}_q&=\qquad\qquad R_{qf}K\hat{e}\qquad\qquad\qquad+R_{qt}\hat{t},\end{aligned}\right\}\quad\ldots\ldots(34)$$

where $\hat{e}$ and $\hat{d}$ are group factors, and $\hat{y}$ and $\hat{t}$ are specifics. All these factors are uncorrelated and normalized.

We shall now derive an expression for V_{ww} in terms of R_{pp}, V_{pp} and R_{ps}, the loadings of the specifics in the first p tests. It will be shown that

$$V_{ww}=(R_{ps}^{-2}+V_{pp}^{-1}-R_{pp}^{-1})^{-1}. \qquad\ldots\ldots(35)$$

By (32) we have $\quad \hat{w}=\hat{z}_p-V_{pp}R_{pp}^{-1}R_{pf}K'^{-1}\hat{e}.$

Hence

$$V_{ww}=\hat{w}\hat{w}'=(\hat{z}_p-V_{pp}R_{pp}^{-1}R_{pf}K^{-1\prime}\hat{e})\,(\hat{z}_p-\hat{e}'K^{-1}R_{fp}R_{pp}^{-1}V_{pp}).$$

Now $\hat{z}_p \mathscr{E}' = \hat{z}_p \hat{f}' K^{-1'} = V_{pf} K^{-1'} = V_{pp} R_{pp}^{-1} R_{pf} K^{-1'}$ (by (26) and (19)). Again,

$$\mathscr{E}\mathscr{E}' = I \quad \text{and} \quad KK' = V_{ff} = I - R_{fp} E_{pp} R_{pf}.$$

After expanding the expression for V_{ww} and using these results we obtain, on making some obvious simplifications,

$$V_{ww} = V_{pp} - V_{pp} R_{pp}^{-1} R_{pf} (I - R_{fp} E_{pp} R_{pf})^{-1} R_{fp} R_{pp}^{-1} V_{pp}.$$

For brevity, we shall write this

$$V_{ww} = V_{pp} - V_{pp} R_{pp}^{-1} T_{pp} R_{pp}^{-1} V_{pp}, \qquad \text{......(36)}$$

where $T_{pp} = R_{pf} (I - R_{fp} E_{pp} R_{pf})^{-1} R_{fp}.$

Writing the second factor of the left member as an infinite geometric progression* we find

$$T_{pp} = R_{pf} \{ \sum_{i=0}^{\infty} (R_{fp} E_{pp} R_{pf})^i \} R_{fp}. \qquad \text{......(37)}$$

We now use the fact (see (30), p. 300) that

$$R_{pf} R_{fp} = R_{pp} - R_{ps}^2 = D_{pp}, \quad \text{say.}$$

The terms on the right-hand side of (37) then become in turn (we are dropping suffixes where no mistake is possible):

$i = 0$: $R_{pf} R_{fp} = 0,$

$i = 1$: $R_{pf} (R_{fp} E_{pp} R_{pf})\, R_{fp} = DED, \ldots,$

$i = 2$: $R_{pf} (R_{fp} E_{pp} R_{pf})\, (R_{fp} E_{pp} R_{pf})\, R_{fp} = DEDED, \ldots,$ etc.

Hence $T_{pp} = D + DED + DEDED + \ldots,$

$$T_{pp} = [I + (DE) + (DE)^2 + \ldots]\, D,$$

and again using the formula for a geometric progression

$$T_{pp} = (I - DE)^{-1} D.$$

Substituting this in (36) we obtain

$$V_{ww} = V_{pp} - V_{pp} R_{pp}^{-1} (I - DE)^{-1} D R_{pp}^{-1} V_{pp},$$

$$V_{ww} = V_{pp} R_{pp}^{-1} (I - DE)^{-1} [(I - DE) R_{pp} - D R_{pp}^{-1} V_{pp}].$$

By (12a) and (38) we find

$$I - DE = I - (R_{pp} - R_{ps}^2)\, (R_{pp}^{-1} - R_{pp}^{-1} V_{pp} R_{pp}^{-1});$$

hence, after expanding and simplifying,

$$I - DE = V_{pp} R_{pp}^{-1} + R_{ps}^2 R_{pp}^{-1} - R_{ps}^2 R_{pp}^{-1} V_{pp} R_{pp}^{-1},$$

---

* It can be shown that the conditions for convergence are fulfilled, i.e. that the modulus of each latent root of $R_{fp} E_{pp} R_{pf}$ is less than unity ((8), p. 98).

and
$$(I-DE)\,R_{pp}-DR_{pp}^{-1}V_{pp}=R_{ps}^{2},$$
whence
$$V_{ww}=V_{pp}R_{pp}^{-1}(V_{pp}R_{pp}^{-1}+R_{ps}^{2}R_{pp}^{-1}-R_{ps}^{2}R_{pp}^{-1}V_{pp}R_{pp}^{-1})\,R_{ps}^{2},$$
$$V_{ww}=\{R_{ps}^{-2}(V_{pp}R_{pp}^{-1}+R_{ps}^{2}R_{pp}^{-1}-R_{ps}^{2}R_{pp}^{-1}V_{pp}R_{pp}^{-1})\,R_{pp}V_{pp}^{-1}\}^{-1}.$$
After multiplying out the factors in the curled bracket we finally get
$$V_{ww}=(R_{ps}^{-2}+V_{pp}^{-1}-R_{pp}^{-1})^{-1}. \qquad \text{......(38)}$$

Thus formula (35) is established, and the theorem stated on p. 297 is proved. Our results can be summarized by saying that the new factorial matrix of the battery of tests is given by
$$\hat{F}=\begin{bmatrix} V_{pp}R_{pp}^{-1}R_{pf}K^{-1\prime}A_{pd} & \vdots\; B_{pv} \\ R_{qf}K & R_{qt} \end{bmatrix}, \qquad \text{......(39)}$$
where B_{pv} and R_{qt} are diagonal matrices containing the loadings of the specifics. The matrix K is defined in (14) and the matrices A_{pd} and B_{pv} are obtained by factorizing the $p\times p$ correlation matrix V_{ww} (35) into a minimum number l of group factors and p specifics, thus
$$V_{ww}=[A_{wd}\mid B_{wd}]\,[A_{wd}\mid B_{wd}]'=A_{wd}\,(A_{wd})'+B_{wd}^{2}. \qquad \text{......(40)}$$

VII. THE NUMBER OF INDUCED GROUP FACTORS

As was shown in the preceding section, the question as to whether or not group factors will be created by selection, depends entirely on the nature of the matrix V_{ww}. Using a result which is due to Thurstone ((7), p. 76) we can generally say that the number l of induced group factors will, at any rate, satisfy the inequality
$$l\leqslant\left\{\frac{(2p+1)-\sqrt{(8p+1)}}{2}\right\}, \qquad \text{......(41)}$$
where $\{x\}$ denotes the least integer equal to, or above x. For example, if p is equal to 2 or 3, at most one additional group factor will appear. On the other hand, in certain cases, l may remain well below the limit set by (41) and we shall now investigate *under what conditions the number* l *may be zero*. This means that the minimum attainable rank of V_{ww} must be nil; i.e. that V_{ww} must be a diagonal matrix. But if V_{ww} is a diagonal matrix, so is
$$V_{ww}^{-1}=R_{ps}^{-2}+V_{pp}^{-1}-R_{pp}^{-1}\quad\text{(by (35))}.$$

Again, since R_{ps}^{-2} is a diagonal matrix, it follows that

$$V_{pp}^{-1}-R_{pp}^{-1}$$

must be a diagonal matrix. We have therefore proved the criterion:

Selection in p *variables will not give rise to additional group factors, if and only if the matrices* R_{pp}^{-1} *and* V_{pp}^{-1} *are identical, except for the diagonal cells.*

VIII. APPLICATION TO THE THEORY OF UNIVARIATE SELECTION

The results that were arrived at in the two articles on univariate selection (1, 2) can, of course, be deduced as special cases from the above theory of multivariate selection ($p=1$).

As an example we shall here obtain the value of the communality h_1^2 of the directly selected variable. In the papers referred to it was found (p. 455, formula (4)) that

$$h_1^2=\frac{H_1^2(1-q_1^2)}{1-H_1^2q_1^2}, \qquad \ldots\ldots(42)$$

where H_1^2 is the communality of the first test before selection and

$$p_1^2=1-q_1^2 \qquad \ldots\ldots(43)$$

is the new variance.

In the notation of the present article we have

$$R_{pp}=[1], \quad R_{ps}=[u_1], \quad V_{pp}=[p_1^2], \quad B_{wv}=[b_1], \quad \ldots\ldots(44)$$

all these matrices being simply scalar quantities. Since u_1 is the loading of the specific in the first test, the communality (before selection) is given by

$$H_1^2=1-u_1^2. \qquad \ldots\ldots(45)$$

Now from the first equation (34) we see that b_1 is the loading of the specific after selection, whence, on standardizing, we get

$$h_1^2=\frac{p_1^2-b_1^2}{p_1^2}.$$

Putting $p=1$ in the relation (41) we infer that $l=0$; i.e. *no new group factor appears in univariate selection*, and (40) simplifies to

$$V_{ww}=B_{wv}^2,$$

or, on using (35) and (44), to

$$(u_1^{-2}+p_1^{-2}-1)^{-1}=b_1^2.$$

On substituting for b_1^2 we get

$$h_1^2 = 1 - p_1^2(u_1^{-2} + p_1^{-2} - 1)^{-1},$$

which, in virtue of (43) and (45), is readily seen to be equivalent to

$$h_1^2 = \frac{H_1^2(1-q_1^2)}{1-H_1^2 q_1^2}.$$

This proves (42). The communalities of the other variates and the new values of the correlation coefficients could be deduced in a similar manner.

REFERENCES

(1) Thomson, Godfrey H. (1938). This *Journal*, xxviii, 451–9.
(2) Ledermann, Walter (1938). This *Journal*, xxix, 69–73.
(3) Thomson, Godfrey H. (1920). *Psychol. Rev.* xxvii, 173–90.
(4) Brown, Wm. & Thomson, Godfrey H. (1921 and 1925). *The Essentials of Mental Measurement.* Camb. Univ. Press.
(5) Aitken, A. C. (1934). *Proc. Edinb. Math. Soc.* (2) iv, 106–10.
(6) —— (1936). *Proc. Edinb. Math. Soc.* (2) v, 37–40.
(7) Thurstone, L. L. (1935). *The Vectors of Mind.* Chicago.
(8) MacDuffee, C. C. (1933). *The Theory of Matrices.* Berlin.

(*Manuscript received* 3 *November* 1937)

EDITOR'S NOTE

Correction: For Blackie, this *Journal*, xxix, p. 187, read T. P. Black *and* John Mackie

PUBLICATIONS RECENTLY RECEIVED

A Source Book of Gestalt *Psychology*. Prepared by WILLIS D. ELLIS. With an Introduction by Prof. K. KOFFKA. London: Kegan Paul, Trench, Trubner and Co., Ltd. 1938. Pp. xiv + 403. 21*s*. net.

Dr Ellis offers his *Source Book* in the hope that it may supplement Prof. Koffka's *Principles* in effecting "the necessary transition from the German psychology of 1912 to the present day". He has prepared some thirty-five abridged transcriptions of studies originally published in Germany between 1915 and 1929 by the major prophets and protagonists of the *Gestalttheorie*. From this book, Dr Ellis suggests in the Preface, the reader may hope to gain "some knowledge of what underlies the *Gestalt* theory both in fundamental findings and, *above all*, in the manner of thinking, of addressing oneself to problems, and in the utilisation of results".

The selections comprising the *Source Book* are divided into three principal groups. The first section (General Problems) contains abstracts of three articles of general theoretical interest, of which two are by Wertheimer and one by Köhler, together with selected passages from the somewhat esoteric and hitherto untranslated *Physische Gestalten*. The second section (Special Problems) is primarily devoted to experimental applications of the *Gestalt* principle. Needless to relate, the most favoured topic is visual perception, and no fewer than nineteen out of the twenty-one strictly experimental papers deal with investigation in this field. Wertheimer is represented here by his pioneer exposition of the conditions of bi-dimensional organization, Fuchs by his interesting studies of transparency and assimilation, Gelb by a condensation of his well-known article on colour constancy, and Benary by his ingenious and important study of the influence of form on brightness contrast. The field of stroboscopic movement is fittingly honoured by the inclusion of four selections, providing serviceable summaries of the explorations of Ternus, Hartmann, Lindemann and Duncker. In the domain of animal psychology, we have abstracts of Köhler's original experiments on relational choice and of Hertz's work upon configurational perception in insects and birds. Outside the field of perception, however, the experimental material is admittedly scanty. Indications of the applicability of certain *Gestalt* conceptions to memory are nevertheless suggested by the inclusion of selections from the studies of Zeigarnik, Gottschaldt and Wulf. Dr Ellis must also be congratulated upon his wisdom in including five selections from the *Psychologische Analyse hirnpathologischer Fälle*. Much of this work of Gelb, Goldstein and Fuchs is obviously of first-rate importance, and it has hitherto been accorded far too little recognition by the ordinary laboratory psychologist. One may hazard the opinion that Fuch's observations upon completion phenomena in hemianopsia furnish by far the most convincing illustrations of configurational tendencies in visual organization.

A few primarily theoretical articles are also included in the second section. Of these, the most interesting are perhaps the two papers of Wertheimer on the thought processes. The final group of selections is exclusively composed of replies to the criticisms of contemporary workers. Koffka politely ferrets out inconsistencies in Benussi's Production Theory, whilst Köhler, in grimmer mood, parries the onslaughts of Rignano and G. E. Muller.

Dr Ellis deserves warm congratulation upon the competence with which he has carried through an ambitious and exacting project. Yet the outcome of his labour is open to serious criticism. Might not the compiler of this volume have performed a more valuable service had his initial aspirations been altogether less exalted? The *Source Book* irresistibly suggests a kind of family album of the *Gestalt* group, and it is tempting to picture those who have lived and worked in its tradition turning the pages slowly and sentimentally, whilst images of the Berlin *Psychologisches Institut*, before the era of *Gleichschaltung*, rise sadly and lovingly to mind. For the research worker to-day, a rather more extensive collection of rather less abbreviated transcriptions would have possessed unparalleled value, and the omission of some, at least, of the theoretical papers would have made possible the inclusion of important recent work, such as that of Metzger and von Restorff. Alternatively, a volume of unabridged translations of the early and relatively inacessible theoretical papers of Wertheimer, Kohler and Koffka would have received an enthusiastic welcome from those whose principal interests lie in the development of psychological theory. Taking the book as it is, however, the choice of material cannot altogether escape criticism. For instance, Hornbostel's whimsy, *The Unity of the Senses*, has long been available in a complete English translation, and it is difficult to believe that there can be reasons other than sentimental responsible for its inclusion in the present volume. Von Frey's essay on tactual sensibility is exceedingly slight, and one cannot avoid the suspicion that its presence here depends largely on the sympathy shown by this distinguished physiologist towards certain of the notions embodied in the *Gestalt* doctrine. Many will deplore the altogether inadequate representation of Koffka; his early work on perceived motion would surely appear more worthy of transcription than the rather questionable researches on the γ movement included in the present volume.

On the other hand, however, all serious students will welcome the opportunity of reading in English so much of the work of Wertheimer. Perhaps this book will emphasize more than any other the enormous debt which psychology, and not merely *Gestalt* Psychology, owes to this highly original and significant thinker. Dr Ellis deserves our sincere appreciation, not only for his translations of Wertheimer, but also for making available (particularly to those whose knowledge of German is severely taxed by the length and complexity of the average *Forschung* article) such serviceable abstracts of so many important experimental studies. His translations are uniformly competent; in one selection only (Lewin) is criticism likely to focus directly upon the translation.

The final section of the *Source Book* will probably alienate a not inconsiderable number of readers. Surely the publication of retorts to criticisms without the merest abstract of the original argument smacks strongly of the totalitarian attitude and is highly undesirable.

Primary Mental Abilities. By L. L. THURSTONE. University of Chicago Press; Cambridge University Press. 1938. Pp. x+121. 9*s*. net.

This is the first of a series of Psychometric Monographs issued by the Psychometric Society. The author of this monograph described, in *The Vectors of Mind* (1935), a method of factorial analysis which he has subsequently applied to a large body of experimental data, and has presented the results in this book. Without familiarity with the earlier work the reader would find difficulty in following certain sections in the present monograph, but if he is willing to accept without question the validity of the method of factorial analysis used, he will find the experiment and results obtaine described in full, clear detail.

The subjects of the experiment were mostly university students, numbering 240 in all, with a modal age of 18; and the fact that their average intelligence (as measured by the psychological examination of the American Council on Education) was exceptionally high shows them to form a highly selected group. They were given a battery of fifty-seven tests, "assembled so as to represent a fairly wide range of the mental activities that are typical in current psychological tests, with special emphasis on those tests which are used as measures of general intelligence". The tests were tentatively grouped under the headings: Abstraction, Verbal Tests, Space, Form, Number, Numerical Reasoning, Verbal Reasoning, Space Reasoning, Rote Learning and an Unclassified group which included tests of Rhythm, Sound Grouping, Vocabulary, etc. These tests took fifteen hours to give and were spread over five days. The interest and attention of the subjects over this long period were ensured by using subjects who volunteered for the purpose and by promising them each an individual vocational report on the test findings. A description and sample of each test is given in the monograph, which is very useful to the reader who wishes to examine critically the factors isolated in the subsequent analysis.

The first step in the factorial analysis of the data obtained from these tests was the computation of the 1596 coefficients of correlation between the test scores. This huge task was substantially decreased by calculating tetrachoric coefficients instead of the more usual product-moment coefficients. In justification of this the author states: "The most complete procedure would seem to be to normalize each of the distributions of raw scores and then to compute the product-moment coefficients. But this assumption that the given distributions of raw scores require normalizing before entering the correlation tables is also just the assumption that is implied by the tetrachoric coefficient. Hence, in using tetrachoric coefficients, we are estimating the product-moment coefficient for the normalized distributions of scores" (p. 58). Doubts may arise in the reader's mind about the validity of this substitution;[1] but this review is not the place for attempting a criticism of the technique of factorial analysis and the assumptions underlying it, so that beyond a caveat to the reader that he should satisfy himself that the methods used are legitimate, no further remarks about the statistical procedures used will be made.

The remainder of the analysis is completed according to the method explained in the author's earlier work and the following factors are isolated:

(1) *The S factor.* This factor is visual or spatial in character and appears to be a common element in thirteen of the tests used.

(2) *The P factor.* This is a perceptual factor of a particular type, i.e. the perceptual function here seems to be a facility in perceiving detail that is imbedded in irrelevant material.

(3) *The N factor.* A numerical factor similar to that identified by Spearman and Kelley.

(4) *The V factor.* This is a factor which can be called "verbal relations", with the reservation that still another verbal factor involves verbal material in a psychologically different manner.

(5) *The W factor.* The other verbal factor, which seems to have as its principal characteristic a fluency in dealing with words, and which is separate from the *V* factor, which is concerned with ideas and meanings.

(6) *The M factor.* The factor of memory.

(7) *The I factor.* An inductive factor, characteristic of tests which demand that the subject shall find a rule or principle for each item in the test.

There remain two factors interpretation of which is tentative and not so clear as that of the preceding seven. These are described by the author as:

(8) *The R factor.* The common characteristic of the tests which exhibited this factor was that the successful completion of the tests involved some form of restriction in the solution.

(9) *The D factor.* A deductive factor entering into four tests only.

The author concludes by a few illustrations of how the 'mental profiles' of certain subjects of the investigation may be related to the vocational wishes of those subjects.

The monograph shows a sincere and painstaking attempt to exemplify the value of the author's method of analysing the factors involved in test performance, and the results are interesting and suggestive to psychologists. It must be borne in mind, however, that the isolation and naming of a factor by any method of factorial analysis does nothing further than prove the existence of this factor in the particular set of test scores under survey, and does not necessarily mean that there is an exact psychological counterpart in the subjects who originally performed the tests. The performance of the same tests by other groups of subjects might quite conceivably yield different results. In this case in particular, a highly selected group of subjects was tested, and it would be of great interest to see if the results were repeated, and the same factors isolated, if an unselected group of subjects (among whom the test intercorrelations would almost certainly be much larger) were given the same tests. Nor does the isolation and naming of a factor prove that a mental ability corresponding to it is possessed in fixed quantities by the subjects tested (and, arguing from them, by mankind in general), so that we can say of a certain person that he possesses abilities $A, B, C, D, \ldots$ to certain specified degrees. We are far yet from being able to do this and it may even be argued that envisaging such a procedure is undesirable, since it is suggestive of a retrogressive step in the history of psychology. E. G. C.

---

[1] Cf. *An Introduction to the Theory of Statistics*, by Yule and Kendall, 11th ed., p. 253: "The product-sum correlation and the tetrachoric correlation are two entirely different measures with quite different properties. The one is in no sense an approximation to the other, and the two may differ largely."

Statistical Tables for Biological, Agricultural and Medical Research. By R. A. FISHER and F. YATES. London: Oliver and Boyd. 1938. Pp. 90. 12*s*. 6*d*. net.

This is a useful set of tables which should be of service to psychologists. It includes the well-known tables of Fisher's *Statistical Methods for Research Workers* (normal deviates, t for small samples, χ^2, etc.), usefully extended by the inclusion of a new column showing values for $P=0{\cdot}001$. The tables for use in the analysis of variance have been much extended by the inclusion of more values of P and also by the provision of a table from which P can be calculated direct from variance ratios without taking logarithms. Of particular value to psychologists should be Table XX for replacing ordinals in ranked data by standard scores. There are also tables of factorials up to 300 and of logs, squares, etc. Although these latter are easily available in less expensive tables, it considerably enhances ease of working to have them, as here, conveniently spaced on a large page.

Philippine Pagans: The Autobiographies of Three Ifugaos. By R. F. BARTON. London: Geo. Routledge and Sons, Ltd. 1938. Pp. xvii + 271. 15*s*. net.

This book inevitably challenges comparison with another recent set of autobiographies of native people, Marjory Perham's *Ten Africans*. From such a comparison it emerges badly. It is said that the three selected Ifugao natives, two men and one woman, were "allowed to tell their stories with a minimum of questioning", having been merely "asked to tell the things in their lives that they considered most important". But the author's interpolated and bracketed remarks seem to make it fairly clear that, wittingly or unwittingly, they were being pushed all the time in the direction of a rather detailed account of their sexual adventures. The book does not, in fact, give the impression of presenting any full and rounded account of life in any community, and although the biographies presented are undeniably lively in various ways, they have the appearance of a good deal of witting or unwitting influence of selection. On the other hand, the method, very briefly indicated and illustrated towards the end of the book, of asking native informants how they would solve certain social problems given them by the enquirer, is of considerable interest and might be extended and exploited to great advantage in other regions.

Reaction to Conquest: Effects of Contact with Europeans on the Pondo of South Africa. By MONICA HUNTER. London: International Institute of African Languages and Cultures (Oxford University Press). 1936. Pp. xx + 582. 30*s*. net.

Dr Monica Hunter has written an extremely comprehensive, absorbing and valuable study, of the utmost importance for the social psychologist who is looking for actual evidence concerning the problems of the contact of social groups. Both by training and by sympathy she has a very thorough knowledge of the Pondo people, and she writes with evident authority and sincerity. She deals with Pondoland itself, with a representative urban community and with Bantu on European farms. She passes in detailed review most of the elements of their culture which have been definitely affected by contact with Europeans. She shows how different may be the effects of such contact as between individual members of either side concerned and as between more or less organized and official groups. She is throughout remarkably sane, sensible and balanced, and she manages to be all these without any sacrifice of vitality or interest. General Smuts writes an introduction to the book in which he says: "To me, who have been in contact with the native mind all my life, the book is a fascinating account of a native tribe, full of curious facts and sidelights, not only on the native mind but on human nature as a whole." The comment is justified. The volume is not merely one more monograph on a native people; it is a notable contribution to an understanding of some of the most important problems in current sociology and social psychology.

Personality: A Psychological Interpretation. By GORDON ALLPORT. London: Constable and Co. 1938. Pp. xiv+588. 16*s*. net.

This is the London edition of a book already reviewed in *this Journal* (XXVIII, 1938, 472). It differs in no respect from the American publication, but the work is a very important contribution to general and to experimental psychology, and many English readers will be glad to know that it is now readily accessible.

Mental Conflicts and Personality. By MANDEL SHERMAN. London: Longmans and Co. 1938. Pp. viii+319. 12*s*. 6*d*. net.

It is urged that all major and most minor difficulties of personality are due to mental conflicts. At the same time it is an error to conclude from this that all conflicts are bad, in the sense that they must lead to diminished efficiency. Many are inevitable, and the real problems arise, not from the conflicts themselves, but from the way they are faced and the solutions that may be attempted. Dr Sherman deals first with the nature and genesis of conflicts, and then proceeds to illustrate and discuss the common types of conflict as they are to be found in contemporary individual and social behaviour. He considers culture conflicts, conflicts of inferiority and insecurity, sex conflicts, conflicts and neuroses, and conflicts and antisocial behaviour. As is inevitable in a short book covering a great amount of ground, a good many special topics hardly receive the attention they deserve. As a sensible introductory study, however, the book has much for which it can be recommended. It cannot be said to rise to any great heights of originality, but it is clear, objective in outlook, the case material is new and mostly very well selected, and the book as a whole will leave the intelligent student with a very good picture of the fields of study with which it is concerned.

A Pediatrician in Search of Mental Hygiene. By BRONSON CROTHERS. London: Oxford University Press (Humphrey Milford). 1937. Pp. xix+271. 8*s*. 6*d*. net.

Dr Crothers describes himself as "a conventionally trained pediatrician who has kept up an interest in neurology". He is convinced that "it is impossible to deal wisely with children handicapped by disorders of the nervous system unless the educational and emotional elements of each situation are carefully considered. Both as a practitioner and as a teacher of medical students I have recognized with increasing clearness that pediatricians are, as a rule, quite unaware of many of the stimulating and highly significant contributions of the psychiatrists, the psychologists, the teachers and the social workers in regard to children". The book is an attempt to argue the case for more psychology for the medical student, and a closer co-operation between the doctor and the psychologist and psychiatrist in the treatment of disease. It is practical in that it not only states the case, but shows how the co-operation regarded as necessary could be, and in some instances, has actually been achieved. The argument is sober and entirely free from exaggerated statement. Dr Crothers recognizes faults and failings on both sides and wisely does not refrain from stating these. It is an American book and the English reader will be particularly interested in the opportunity which it gives him to compare conditions in that country with those which obtain in his own. There appears to be less difference than is often supposed. Some of the details of teaching and hospital administration which are advocated would need alteration before they could well be adapted to the English situation, but the general argument is closely applicable here also, and it is greatly to be hoped that the aims which Dr Crothers states clearly and concisely may also direct the future development of pediatrics in England.

Crooked Personalities in Childhood and After. By R. B. CATTELL. London: Nisbet and Co., Ltd.; Cambridge University Press. 1938. Pp. xi + 215. 7*s*. 6*d*. net.

This somewhat unpromising title is given to what is, nevertheless, a very useful and readable introduction to the psychotherapy, mainly of children, but with reference also to adult adjustment. The author gives brief outlines of the theories of Freud, Jung and Adler, as they relate to clinical practice, and describes the Freudian 'mechanisms' in more detail. It is no doubt of value to attempt to relate the findings of theoretical and experimental psychologists to those of psychopathologists, and apply the results of the former to clinical diagnosis. But Dr Cattell's attempt is too superficial and uncritical to be helpful—in contrast with his treatment of his psychopathological material. The chapters on the Work of the Clinic and the Family Background form the best part of the book; and the relation of maladjustment to the life and values of society as a whole, especially with regard to the maladjustments of genius, suggests a useful corrective to the pursuit of 'normality' at any price.

Both the student and the layman will find this book most valuable in tackling the involvements of psychopathology, particularly as it emphasizes the importance of a thorough study of the individual circumstances in each case, rather than of adherence to any pre-conceived theory.

The Ego and the Mechanisms of Defence. By ANNA FREUD; translated by C. BAINES. International Psycho-analytical Library, No. 30. London: The Hogarth Press and the Institute of Psycho-analysis. 1937. Pp. x + 196. 9*s*. net.

There has been a gradual tendency in psycho-analytical circles to widen the scope of analysis, to include that of the ego and super-ego as well as that of the id, and thus to bring psycho-analytical practice more into line with that of non-psycho-analytical practising psychologists.

In this book the author deals with the non-sexual factors and discusses the defence mechanisms, both normal and pathological, employed by the ego in its struggle with the id and super-ego, and the consequent effects on behaviour. She shows that, from a psycho-analytical standpoint, it is not to be expected that the mere removal of restraining influences in early childhood will prevent the development of neurosis which is a consequence of the ego's fear of the strength of the instinctive forces in the id. She considers at length the many character changes which occur before and at puberty; and discusses the asceticism and intellectualism of the adolescent. The exposition has that clarity one associates with Anna Freud, and the ideas—whether or no one agrees with the explanations—will be of interest, not only to analysts, but also to all who have to do with children.

Peter Kurten: A Study in Sadism. By GEORGE GODWIN. London: The Acorn Press. 1938. Pp. 58. 5*s*. net.

This book is to be taken as a foreword to the same author's translation of Prof. Karl Berg's full-dress study of the same topic which it is hoped may appear under the title of *The Sadist*. It is a brief, unvarnished study of this man who appears to have been a fairly reasonable, mild-mannered, even attractive person in everyday life, but who, in his night exploits of brutal murder and sexual outrage, was an uncontrolled monster of depravity. The book is, of course, for the specialist; but it contains a few points of general psychological interest. For example, why should Kurten have had the most detailed memory for every event of his most disgusting crimes? Why should he, even in ordinary life, have felt regret for his infidelities, and none for his brutalities? Is it true to maintain, as the author, following Prof. Berg, does maintain, that Kurten's crimes might have been committed by Baudelaire, by Verlaine, or even by Goethe, except that all these three had achieved a catharsis of words? Have words such power? Perhaps the fuller study may provide convincing answers.

They Always Come Back. By CICELY MCCALL. London: Methuen and Co. 1938. Pp. xxiv+258. 7*s*. 6*d*. net.

It is no use writing an ordinary descriptive review of this book. The volume must be read. It is an account of first-hand experiences as an officer at Holloway and at the girls' Borstal institution at Aylesbury. Miss McCall is, however, no ordinary officer. She has psychological training and, what is much more to the point, she evidently possesses also genuine psychological insight. She tells a perfectly plain story of life in the prison and at the institution and tells it in an engaging personal manner, writing all the time, not about cases, but about people. There is no sentimentalizing and no high-flown moralizing. Undoubtedly the record is selected as any such record must be, but it seems utterly impossible that anybody can read this book—and it is greatly to be hoped that a vast number of people will read it—without a vivid realization that there is a great deal in the prison and Borstal systems of this country that is psychologically wrong. More than that, the reader will be left with clear ideas of *what* is psychologically wrong and of how a large amount of it could quite easily be set right. The book is certainly out for reform, but it is written throughout with exceeding good humour. Its influence ought to be wide and beneficent.

Play in the Infants' School. By E. R. BOYCE. London: Methuen and Co. 1938. Pp. xiii+188. 5*s*. net.

In 1933 Miss E. R. Boyce was appointed head teacher of the Infants Department of the Raleigh School, Stepney, London, E. She found 300 children on the roll, with an age range from three to seven-and-a-half years. They were all drawn from a poverty stricken area, many of them under-nourished, physically and mentally retarded, restricted and none too happy. She decided that the school must be reorganized, the children given more freedom, their school tasks made more of the order of a joyful adventure than of mere duties, and their physical development as much considered as their conventional mental training. Here is the record of the experiment which she undertook. It carries the story up to 1936. It gives full detail of equipment, time-tables, methods, difficulties and successes. As Dr Susan Isaacs says in a brief

Foreword: "Miss Boyce has shirked none of the difficulties of her enterprise. She shares with the readers her doubts and hesitation, her gropings and her failures, as well as her successes. She assesses the achievements of the school in ordinary scholastic terms, but is able to do more than this: she shows how for these children the level of actual attainment is far less significant than the fostering of the will to learn and the belief that learning has some value for the personal life." The story is fascinating and is well-told. Now it is greatly to be hoped that at least some of these children will be followed out of their Infants' school into their Primary and Secondary school life, so that it may become clear to what extent the lessons of adjustment and enthusiasm which seem to have been very successfully learned in the early stages persist in the larger world. There certainly appears every reason to hope and believe that they may do so.

The Assistant Master Speaks. Twelve Essays by Assistant and House Masters at Public and Secondary Schools. 1938. London: Kegan Paul, Trench, Trubner and Co., Ltd. Pp. viii+296. 7*s.* 6*d.* net.

These essays were obtained by inviting contributions from all Assistant Masters in Secondary Schools. It is not stated who made the selection here given; and the twelve essays are of very unequal merit. Two good constructive contributions on the improvement of the curriculum in Secondary Schools are made by C. D'O. Gowan and G. F. Golding, of Eton College and St Olave's Grammar School, London, respectively; and there is an excellent essay on Discipline by a House Master at an English Public School. These are full of sound psychology and practical wisdom. The other essays in the main reflect the dissatisfaction of school masters and of the general public with the present system of secondary education for boys, and make certain suggestions for remedying this state of affairs. But with one or two exceptions (notably that of Donald Hughes, of the Leys School) they are lacking in conviction.

An Introduction to a Christian Psycho-Therapy. By J. A. C. MURRAY. Edinburgh: T. and T. Clark. 1938. Pp. xii+279. 7*s.* 6*d.* net.

The book is a sincere and well-written argument, influenced in the main by a somewhat eclectic attitude towards modern psycho-therapeutic practice and theory. It is less a statement of the case for grafting psycho-therapy on to Christian thought and practice as for treating psycho-therapy as a natural outgrowth of that part of the Christian religion which is directly concerned with the life of people in the actual world. It is bound to make certain assumptions, one of which, for instance, is the old one of a religious instinct, but if these are granted the treatment runs smoothly and clearly and is supported by many references to original Christian teaching. The author considers that the time is ripe for the training of specially selected candidates for ordination in academic, experimental and especially clinical psychology. There can be little doubt that, if Christian beliefs and psycho-therapeutic practice can be happily combined, in many instances the strength and practical success of both may be greatly increased. But whether it is theoretically necessary, or even possible, for psychotherapy to have a fundamentally Christian basis is much more disputable. Meanwhile the book may have a good influence provided it does not set off inadequately equipped persons in the immediate practice of the techniques which it describes.

BRITISH PSYCHOLOGICAL SOCIETY

REVENUE ACCOUNT AND BALANCE SHEET AS AT 30 SEPTEMBER 1938

Revenue Account for the year ended 30 September 1938

1937 £ s. d.		EXPENDITURE	£ s. d.	£ s. d.
170 0 0		*To* Rent, 55 Russell Square		170 0 0
252 0 0		Secretary's Salary		252 0 0
79 9 5		Printing and Duplicating		93 5 0
32 2 8		Hire of Halls		25 15 6
		Office and General Expenses:		
	£30 5 8	Caretaker, Lighting and Heating	33 6 5	
	10 10 11	Stationery	15 12 1	
	70 0 9	Postages and Telephone	68 0 9	
	4 12 0	Fire, etc. Insurance	4 12 0	
	5 5 0	Subscriptions to outside Associations	4 14 6	
	10 10 0	Audit Fee 1937	10 10 0	
		Renovating Typewriter	6 1 0	
	1 18 1	Repairs	1 12 11	
	12 10 1	Teas	9 15 3	
	8 3 7	Sundry Expenses	9 8 6	
153 16 1				163 13 5
		Journals Publication Expenses:		
	166 9 3	General	263 19 6	
	121 9 4	Medical	124 14 9	
	89 5 10	Educational	100 0 10	
377 4 5				488 15 1
22 3 8		Extended General Meeting 1938 ...		11 16 8
4 5 1		Work for Committee on Human Mental Measurements		19 1
17 5		Midland Branch Expenses		4 4 0
2 12 1		Northern Branch Expenses ...		3 4 6
9 12 5		Scottish Branch Expenses		11 0 5
18 14 3		*Revision of Members' Register*		
1 0 2		Cheque Books and Bank Charges ...		1 0 2
		Depreciation:		
		Library	64 0 7	
84 10 8		Office Furniture and Fittings ...	19 2 9	83 3 4
1208 8 4				1308 17 2
154 14 2		Balance, being Surplus of Income over Expenditure		140 11 0
£1363 2 6				£1449 8 2

1937 £ s. d.	INCOME	£ s. d.	£ s. d.
	By Subscriptions 1938	1424 15 6	
	Subscriptions 1937	10 2 6	
		1434 18 0	
	Less Resignations and Removals ...	109 7 6	
1242 5 4			1325 10 6
120 17 2	Interest on Investments		123 17 8
£1363 2 6			£1449 8 2

[P.T.O.

REVENUE ACCOUNT AND BALANCE SHEET (*continued*)

Balance Sheet as at 30 September 1938

1937 £ s. d.	LIABILITIES	£ s. d.	£ s. d.
1 1 0	Subscriptions paid in Advance		2 2 6
	Publications Reserve:		
122 0 0	Balance as per last Balance Sheet ...		122 0 0
	Revenue Account		
	Balance as per last Balance Sheet ...	4765 14 1	
	Add Surplus of year	140 11 0	
4765 14 1			4906 5 1
£4888 15 1			£5030 7 7

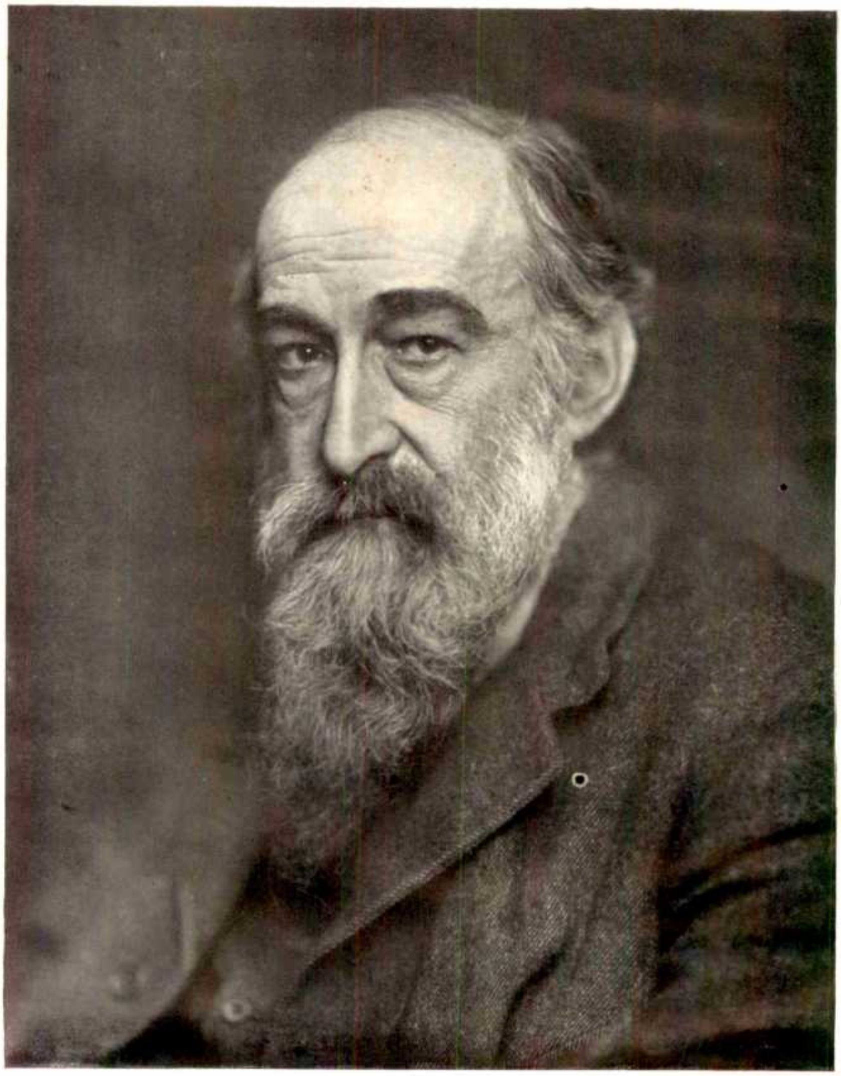
1937 £ s. d.	ASSETS	£ s. d.	£ s. d.
238 8 6	Cash at Barclays Bank Ltd		210 2 1
1 3 0	Postage stamps in hand		16 6
	Investments at Cost:		
	£300 5% Conversion Loan 1944–64 ...	318 9 9	
	£3320. 14*s.* 8*d.* 3½% Inscribed War Loan	3425 7 9	
3543 17 6			3743 17 6
	Note Market value of above Securities at 30 September 1938, £3720		
1 9 3	*Sundry Debtors*		
111 6 0	Subscriptions in Arrear		129 15 0
	Office Furniture and Fittings.		
	Balance as per last Balance Sheet .	177 15 0	
	Additions during year	13 10 9	
		191 5 9	
	Less Depreciation	19 2 9	
177 15 0			172 3 0
	Library:		
	Balance as per last Balance Sheet ...	583 0 0	
	Add Expenditure during year ...	57 5 7	
		640 5 7	
	Less Depreciation	64 0 7	
583 0 0			576 5 0
	Monograph Publications		
	Balance as per last Balance Sheet ..	181 15 10	
	Less Amounts repaid during year ...	34 7 4	
181 15 10			147 8 6
50 0 0	Shares in the British Journal of Educational Psychology, Ltd., at Cost ...		50 0 0
£4888 15 1			£5030 7 7

I beg to report that I have examined the foregoing Accounts with the Books and vouchers of the British Psychological Society and that I have obtained all the information and explanations I have required. In my opinion the Balance Sheet is properly drawn up so as to exhibit a true and correct view of the state of the Society's affairs according to the best of my information and the explanations given to me and as shewn by the Books of the Society. The Securities appearing under the heading Investments have been verified by production of the Bankers' Certificate.

T. BARRETT,
Incorporated Accountant

S. J. F. PHILPOTT,
Hon. Treasurer.

28 *October* 1938.



PROF. SAMUEL ALEXANDER, 1859–1938

Part 1 Wanting.

VOLUME XXIX APRIL, 1939 PART 4

OBITUARY NOTICE

PROFESSOR SAMUEL ALEXANDER, 1859–1938

Professor Samuel Alexander, who had held the Sir Samuel Hall Chair of Philosophy at Manchester University from 1893 to 1924, died at his home at Withington, Manchester, on 13 September 1938, at the age of seventy-nine.

He was born in Sydney, studied at Wesley College, Melbourne, and Melbourne University, and came in 1878 to Balliol College, Oxford. Here his thinking was influenced, as his earlier work shows, by Green and Nettleship. After graduating he was made a Fellow of Lincoln College, where he remained as philosophy tutor from 1881 to 1893. Here he developed his interest in the empirical aspect of psychology, and enriched Oxford's remarkable collection of legends about this study by a story concerning a queer fellow who strapped people to machines intended to measure emotions, and once bounded into the experimental room with the kindest of faces, covered with ornamentations to induce terror. This, however, was not the emotion produced in the victim.

During these years, under the influence of Green, he wrote *Moral Order and Progress*, which, a critic writes, already showed characteristics of Alexander's mature mind; a great respect for fact, and a love of synthesis. In 1893 he succeeded Adamson at Manchester, which was then Owens College. The evolution of Manchester University owes much to Alexander's unique fusion of simplicity, modesty, good fellowship and business ability. Those were the days of trams and bicycles, when the nearby suburbs housed the University staff. The internal combustion engine had not yet begun its calamitous destruction of culture, of which the centrifugal trek from the cities was an early feature. In such a *milieu* of easy "dropping-in", Alexander's influence was all-pervasive. His imperfect hearing, even at that time, was a handicap, but not entirely, for people delighted to speak directly to him. He ignored idle chatter, and sometimes seemed psychically deaf even to relevant questions, proceeding conversationally with the unhurried dignity of a liner entering dock. Perhaps he was aware of this advantage and, too, of the mental mechanism of projection, for once he opened a public discussion by remarking: "When I said this in London, I think Mr — (a dissident philosopher) did not quite hear some of my objections to his views. To ensure that he shall hear properly to-night, I present him with a copy of my paper."

At first, in Manchester, Alexander published little. Writing was impossible when he was sloughing the skin of idealism which had grown upon him in Oxford. Idealists had argued that to be known is to be dependent on mind. In 1905, Alexander began to counter this with assertions which seemed to be sheer paradox, for the Gifford Lectures and their fuller statement, *Space, Time and Deity*, were still below the horizon. To be known, said he, was to be independent of mind. In exemplifying this assertion, the stately trees in Brunswick Road (for this reason alone they ought to be preserved) must have been utilized almost daily.

Alexander's way of exposition was general and abstract. It consisted of cheques drawn upon the bank of knowledge with a masterly, confident, if illegible hand. He vouchsafed the reader few concrete examples, and these were casually lifted from the conventional philosopher's portable pack of trees, tables and teacups.

Perhaps this way of writing led Alexander farther away from the average reader than either of them ever suspected. It is difficult to believe that he noticed the crudity of early behaviourism. In the exciting years, for psychology, of 1910–14, he seemed almost bored by discussions of Freud. "The stuff that dreams are made on", was, in his view, the stuff that physicists ought to study, though he knew them too well to believe that they ever would. He seemed unmoved by speculations about the mechanism of dreams. Two optical images can be superposed by mirrors; what then is mysterious about Freud's 'Condensation'? Alexander even denied the uniqueness of the relation of knowing. He once said, "The teacup knows the table, but it does not enjoy itself; I know the teacup but I also enjoy myself." That was all the difference, for ordinary metaphysicians, but for the sociable Alexander, what an all!

The contrast between his teaching and that of Külpe at Würzburg or Freud in Vienna would have been a shock to any Continental scholar. Alexander, perhaps urged by his Australian blood, was establishing a lonely frontier post of psychology, and looking away from it across a desolate No-man's Land to the adjacent territories of physics, metaphysics and aesthetics. In logic and ethics he now took almost a perfunctory interest.

To understand him between 1910 and 1914, one would have to comprehend Alexander's Manchester, which, to borrow an American description of another place, was not so much a city as a state of mind. That state of mind was almost dominated by three influences from the Antipodes: Alexander, Rutherford, and Elliot Smith. They had slipped into Manchester's culture pattern as easily and naturally as into its University common room.

Manchester's straight talk, her ferocious contempt for appearances, her unconcealed uninhibited friendliness for people she liked, her gentle cherishing of certain cultural values—having made sure that a few extraverted business men were willing and able to pay for them—her local sport of bubble pricking (or bashing), her proud, even defiant cosmopolitanism; with all these notes this trio could harmonize. Alexander's spectacular appearance and cathedral-organ voice, Rutherford's bucolic bellow, Elliot Smith's illusory shyness: without them Manchester's university life cannot be pictured.

It is as difficult to estimate Alexander's influence upon psychology as it is to limit an account of his influence upon psychologists. The central point of his teaching was that a completely conational psychology was possible, that the alleged cognitive element is non-mental, and that feeling is a qualification of conation. Perception can be resolved into act and external thing. This is not the result of common-sense, nor is it naïve realism. That it happens is not a postulate but a description of the event in its simplest terms. On the different levels of experience, perception, imagination, memory, etc., there corresponds to each form of conation a certain form of non-mental object or cognition. Both the object cognized, and even the organic and kinaesthetic *sensa* in which attention results, are objects contemplated. They are not enjoyed, undergone, lived through, or suffered, as is the mental act.

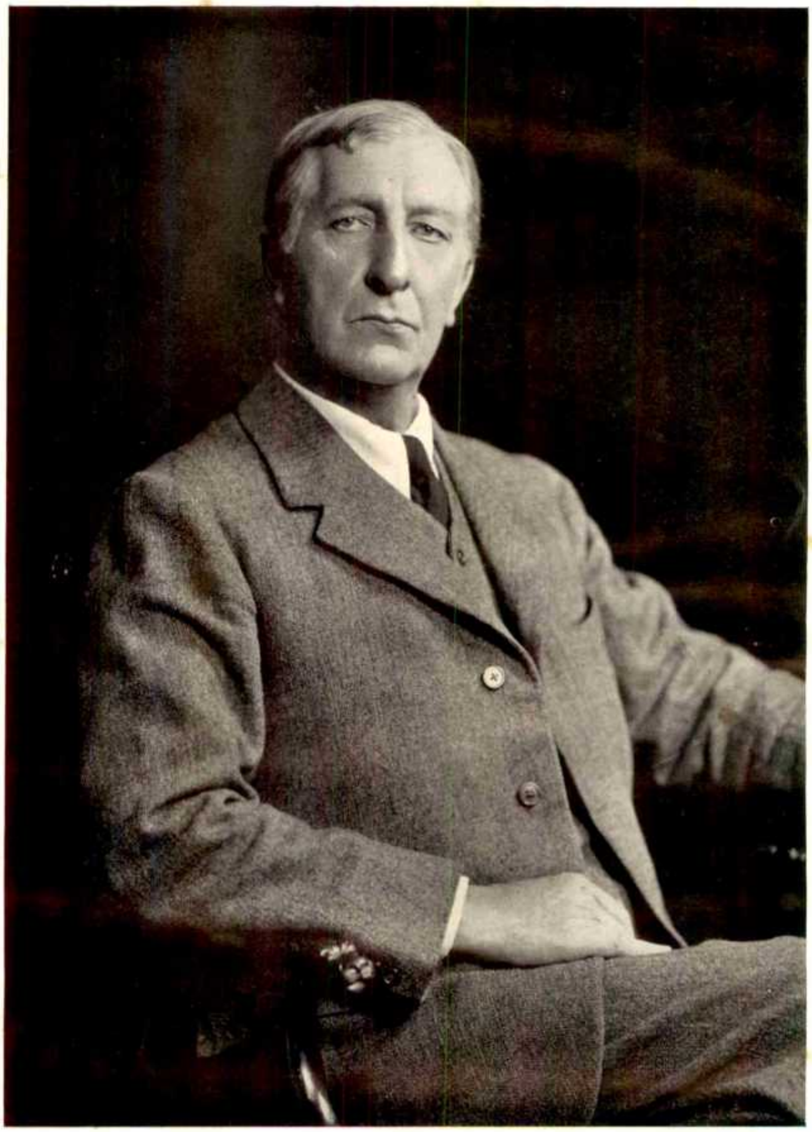
This 'act-object' contrast pervades all his teaching. It is the business of psychology to describe and to analyse enjoyments; this is introspection. The subject, as enjoyed, is the continuous tissue of its acts of conation; consciousness is the general form of such enjoyment. The combination of the subject, who thinks, feels or wills, and the bodily self, he called the person, and distinguished stages in the growth of personality.

Alexander's special interests in acts rather than in objects (in '-ings' rather than in '-eds', to quote Prof. C. Lloyd Morgan) perhaps made him uninterested in those facts which follow the behaviourist around, like neglected animals, with sad, reproachful eyes; synaesthesia and eidetic images, memory images, hallucinations, dreams; indeed, one of Alexander's greatest admirers wrote of his "mistreatment of imagery in memory". Remembering was for him an act of desire directed backwards towards the past. The image, if the original object had been a physical one, is physical also; is the same thing as the percept, appearing in a different form. Deciding that half the problems of the psychologist concerned non-mental happenings, and having handed over the mandate to the unheeding physicists, he busied himself with other matters, and consequently his references to problems of memory seem a little weary and elusive. For the last thirty years of his life Alexander's interests were chiefly in metaphysics, though he wrote actively upon aesthetics, about which the present writer is utterly incompetent to judge. But Alexander was the kindest of critics, and so often used the term 'beautiful', especially of human beings, that his essays on beauty were also essays in benevolence.

If Alexander's character was formed in the southern hemisphere, Oxford moulded his personality, bestowing that combination of sonorous public speaking, impressive diction, and occasional insouciance about its comprehensibility.

Alexander was a mass of contradictions, resolved like the discords in a symphony. Careless of externals, yet amiably consenting to be photographed, drawn, painted, etched and finally sculptured, garbed in a manner which launched a thousand legends, but loving his finery as Public Orator; magnificently declaiming his speeches, composed with the greatest artistry and after patient researches into the graduand's public and private records, yet capable of thanking a beloved friend publicly in a lazy tangle of *clichés* and repeated half-sentences, with naughty beams of affectionate atonement; a lover of all nations, but not a traveller, looking like a cartoonist's dream of a professor, yet punctual and punctilious, unforgettably inspiring, yet producing only a few distinguished students, loving and beloved by man, woman, child and animal, yet lonely, and 'rather liking it'.

T. H. PEAR.



PROF. WILLIAM McDOUGALL, 1871–1938

OBITUARY NOTICE

PROFESSOR WILLIAM McDOUGALL, 1871–1938

WITH the death of William McDougall on 28 November 1938, British psychology lost one of its foremost representatives and the British Psychological Society one of its most distinguished Honorary Members. He was an original member of the Society, the first meeting of which (on 24 October 1901) was, indeed, preceded by a number of informal discussions of a little group which met in his laboratory at University College, London. He acted as Hon. Treasurer to the Society from 1905 to 1908 and again from 1911 to 1916, making during all this time frequent contributions to the Society's meetings. He was an Associate Editor of this *Journal* from its inception, and in 1934 was elected an Honorary Member of the Society, when in that year it was decided—for the first time since the War—that the short list of those whose services to psychology the Society wished to recognize in this way should include representatives of British as well as of foreign nationality.

Born in Lancashire on 22 June 1871, McDougall early showed himself to be the possessor of exceptional abilities, and, as he tells us (in his illuminating contribution to *Psychology in Autobiography*), it was soon made plain to him that he was expected to distinguish himself. After attending a private school, he spent a year at the *Realgymnasium* at Weimar and then, when only fifteen, entered the University of Manchester, 'a serious student', but one whose physical endowments equalled his mental capacities, and who found time to represent his University at Rugby football, on the track and on the river. Four years at Manchester were followed by four at St John's College, Cambridge, and these in turn by yet another four at St Thomas's Hospital, London, towards the end of which period he published his first papers, which were devoted to the subject of muscular contraction. He had no intention, however, of practising medicine, and even as an undergraduate had determined to devote himself to a study of the brain and nervous system, "in which were locked the secrets of human nature", though a reading of James's *Principles of Psychology* during his hospital years revealed to him another and complementary method of approach. After election to a fellowship at his own College at Cambridge, he joined the Cambridge Anthropological Expedition to Torres Straits, then proceeded to Borneo at the invitation of Dr Charles Hose (in collaboration with whom he subsequently produced a large two-volume work on *The Pagan Tribes of Borneo*).

Returning to Europe in 1900, he married Anne Hickmore of Brighton and spent some time at Göttingen with G. E. Müller, where he embarked on the experimental study of vision that occupied much of his time during the next few years. In 1900 he was invited by James Sully to undertake the teaching of experimental psychology at University College, London, where a small laboratory had recently been established (equipped with the apparatus formerly used by Münsterberg at Freiburg), and he continued to teach there till 1906, though in 1904 he entered upon the more important duties of Wilde Reader in Mental Philosophy at Oxford. He was debarred by the conditions of his Readership from dealing in his official lectures with the experimental approach, but Francis Gotch, the Professor of Physiology, put at his disposal, first one, and then (when a new wing was added to the physiological laboratory) six rooms, where, in a private capacity, he was able to continue research, and where he gathered about him a small group of students, including M. E. Bickersteth, William Brown, Cyril Burt, Horace English, A. M. Hocart, May Smith and the present writer. To the members of this group his teaching and companionship were equally delightful. He himself tells us that he regarded them as "brands plucked from the burning", and these words indicate a certain inevitable dissatisfaction with the conditions at Oxford, where psychology had no place in the curriculum, where potential students were discouraged by philosophers in whose opinion psychology could not and therefore did not exist, and where the scientists on their part (with a few exceptions) also regarded with grave suspicion the upstart branch of knowledge which he represented. There were, however, certain compensating advantages, and he made full use of the freedom he enjoyed to lecture on any aspect of psychology he chose, much of his later published work being adumbrated in these early lectures. The informal meetings in the laboratory were, however, for those who were privileged to attend them, a more potent and inspiring influence than the official lectures, although these latter sometimes drew good audiences (especially a course of hypnotism, which was regarded as a very daring subject for a University lecturer to tackle).

McDougall's earlier publications clearly reveal his interest in both the neurological and the psychological approaches to the problems of the mind. Indeed his very first contributions of a psychological nature (three articles in *Mind* of 1898) were directly concerned with problems of psychophysical correlation. There followed a brilliant series of papers, many of them on vision, others on the 'attention process', hypnotism, fatigue, and other subjects, all largely 'psychophysical', in the sense that introspective observations (often based on ingenious experiments) were interpreted to a great extent in physiological terms. His first book, the small *Physiological Psychology* published in 1905, proved a very fascinating general exposition from this point of view and (as the present writer can testify) exercised a great influence on many students in Oxford and elsewhere who were otherwise suspicious of psychology. From this time onwards, however, his interest in neurology and his satisfaction with physiological explanations appear to have waned, giving place to a more biological point of view—his biology, from the start, being dominated by the concept of purposive adaptation. In this change of standpoint we may perhaps trace the conflicting influences of Sherrington and James respectively, both of whom were of great importance in his life, though the inspiration derived from James was destined to have the more permanent effect. Anyhow, from the time of writing his *Body and Mind: A History and Defense of Animism* (1911), McDougall was but little occupied with problems of brain and nerve function. And this book itself stated boldly and clearly the thesis that mind itself must be regarded as a potent cause in development and evolution. By general admission it presents by far the strongest case that has ever been made for Interactionism, and it seems likely to remain a work of supreme importance in this field until further facts of fundamental significance as regards the mind-body relationship shall have been discovered.

Meanwhile, with his second book (the *Introduction to Social Psychology*, 1908) McDougall had entered a new and more purely psychological field, and had definitely emerged as champion of what subsequently became known as the hormic theory, the view that stresses the striving and purposive aspects of all life and that seeks to trace the energy displayed in every activity of living creatures to some inborn disposition or instinct. The book will (as he himself recognized) almost certainly be reckoned his most original contribution to psychology. At any rate, it soon became, and continues even now after more than thirty years, to be one of the most popular books on psychology. Not only did it complete the task (begun by Ward and Stout) of correcting the intellectualist and hedonist bias of the associationist views then dominant both inside and outside the laboratory, but at one stroke it produced an orderly scheme for the understanding of orectic life from its simplest to its most complex manifestations, a scheme which he himself was able to use in all his subsequent writings and which has also been adopted in whole or in part by many other authors.

In the title of this work he had used the term *Introduction* in the rather unusual sense of Prolegomena, and he regarded his study of instinct, emotion, habit, sentiment, etc., as the essential preliminary to a treatment of social psychology itself, which was to constitute his *magnum opus*. The war, however, interrupted his plans, and after the publication of one other small book, *Psychology, the Study of Behaviour* (which also proved extremely popular), he found himself in 1914 driving an ambulance on the Western front, and then—more appropriately—a major in the R.A.M.C., dealing with cases of 'shell-shock'. This work continued till 1919 when, feeling that he needed to digest his material on the functional nervous disorders (which now filled thirty-five notebooks) in the light of the experience of others, he visited C. G. Jung at Zürich and was for a time analysed by him. He came away "enlightened but not convinced", and his book on *Abnormal Psychology* did not appear till 1926, though in the interval he wrote two other works of great importance —*The Group Mind* (1920) and the *Outline of Psychology* (1923). The former was intended to be the first volume of the great projected work on social psychology already mentioned, but he was disappointed at its reception and "the *magnum opus* went a-glimmering", though he subsequently returned to social problems in several minor works; the second was in effect the working up of the principles enunciated in the *Introduction to Social Psychology* into a general system of psychology, into a scheme, as he believed (and as many will agree), "more complete than any other".

Before this, however, McDougall had moved from England to the United States, having in 1920 accepted an invitation from Harvard to occupy the Chair of Psychology left vacant by the death of Münsterberg. As regards prestige, equipment, and number of students, Harvard offered vastly greater opportunities than Oxford; but soon after his arrival McDougall found himself in a behaviouristic atmosphere that was distinctly uncongenial. It is true that he himself had called psychology 'the study of behaviour', but to him behaviour meant conduct determined by purpose and desire, whereas to many of his colleagues and students it meant a materialistic view of life conceived in terms of stimulus-response mechanisms and conditioned reflexes. For this and other reasons, McDougall never fitted quite happily into the American scene. He tells us, however, that he never regretted the adventure, and indeed, as he himself recognized, there was an element in his character which made it impossible for him to feel entirely at home in any group or any environment. Though he was very far from being in all respects a radical, he was endowed with a certain intellectual 'arrogance' combined with a tendency to scepticism, as a result of which he nearly always found himself in opposition to some prevailing tendency or theory. In the early days he combated the fashionable theories of Hering in the field of vision and the even more fashionable solution of the mind-body problem in terms of psycho-physical parallelism, and it was perhaps a tragedy that in one of the few important matters in which he thought he could wholeheartedly adopt the views of another worker (in the case of Shand's doctrine of sentiments), he soon found that there existed irreconcilable differences of opinion which he had not at first suspected. From early years onwards he became increasingly sceptical of the prevailing Weismannism, which he found 'all-dominant' even in his Cambridge days and the influence of which he suspected of being due to its association with 'mechanistic dogma in biology'. In America he resolved to put the opposing Lamarckian theory to the test and began the long series of experiments with white rats that is still in progress—experiments designed to show whether, as Lamarck had supposed and as Weismann had denied, there really existed any tendency for 'acquired characters' to be inherited. The experiment, as is now well known, has produced results pointing clearly in the Lamarckian direction, though their precise significance is still a matter of dispute. Whatever the ultimate outcome, however, the experiment was devised and carried out on a heroic scale and it may indeed yet come to be regarded as of epoch-making significance.

His tendency to oppose prevailing views or theories continued unabated into later life, and indeed many of his last writings were polemical in nature. Behaviourism, psycho-analysis, Gestalt, each in turn became an object of detailed and penetrating criticism. His battle with behaviourism was the longest and bitterest of all, partly because American psychology was so largely influenced by this doctrine in the years following the war, and partly because behaviourism seemed to him to be utterly opposed to his own conceptions of life and mind and based on assumptions which threatened to deprive psychology of all its true significance. In both Gestalt and psycho-analysis there was much that he admired, in particular the new approach to old problems, the ingenious experimentation and the rejection of hedonism and associationism (in the case of Gestalt) and the whole-hearted adoption of the categories of wish or purpose (in the case of psycho-analysis). Moreover McDougall was impressed by the successful application of psycho-analytic concepts to social phenomena and regarded this as one more proof of the correctness and fruitfulness of the hormic view in psychology. But he deplored the vague and, as it seemed to him, utterly inadequate physical conceptions of energy prevalent among Gestalt psychologists; while the psycho-analysts on their part offended him with what appeared to be their inadequate battery of instincts, their overworking of the sex instinct, their hypostatizations, their uncritical resort to psychological hedonism and their somewhat chaotic theorising generally. It is true that much of psychoanalytic theory does appear to be a clumsy makeshift when compared with McDougall's own neat and systematic formulations, though at the same time few will deny that as regards the detailed understanding of individual motives McDougall did not display so deep a penetration as did Freud. Perhaps, in his heroic endeavour to bring order out of chaos, the system he devised was *too* neat and precise to serve as an adequate theoretical framework for the vast complexities of human motivation and behaviour. Indeed McDougall himself has, during the last fifteen years, introduced various extensions and modifications of his original system (e.g. his doctrines of 'temper', 'taste' and the 'derived emotions'), which have given it a less orderly appearance. Nevertheless, some kind of consistent theory was, and still is, an urgent necessity for the fuller understanding of orexis, and McDougall's formulations, whatever further revision they may ultimately require, have beyond all doubt been of immense service to contemporary psychology.

In 1927 McDougall moved from Harvard to Duke University. In his new position he was able to spend the early summer months in England. But whether he was officially in term or in vacation, he was always busily at work, and during the last years books and articles continued to appear, many of them minutely critical, as we have already indicated, others broadly constructive, in the sense of developing the implications and applications of his own purposive view of psychology. Among works of the latter kind were *Character and the Conduct of Life* (1927), *Modern Materialism and Emergent Evolution* (1929), *World Chaos* (1931), *Religion and the Sciences of Life* (1934), *The Frontiers of Psychology* (1934), while *The Energies of Men* (1932) was a restatement of his own position from the point of view of pure psychology. His last work, *The Riddle of Life*, has only just appeared, having been completed after his partial recovery from a severe operation in February 1938.

Apart from what may be regarded as his main activities, McDougall had a number of sidelines in which he also displayed his characteristic energy. He was a keen eugenist and he developed this theme in *National*

Welfare and National Decay (1921) and in several other writings. In *Ethics and Some Modern World Problems* (1924) and in *Janus or the Conquest of War* (1927) he made the interesting suggestion that the national ownership of aircraft (or at any rate fast flying aircraft) should be forbidden, their use being confined to the purposes of an international police force. In view of subsequent events we may perhaps regret that the world did not see its way to treat this proposal with greater seriousness. He also throughout his life displayed an open-minded interest in psychical research, feeling that it was a scandal that psychologists should refuse to lend a hand to the very small body of workers who were really seeking the truth in this much debated field. In this connexion he sponsored the experiments of his colleague, Prof. J. B. Rhine, on Extra-sensory Perception, which have recently aroused such widespread attention.

McDougall was a member of very many scientific bodies and societies, both in Europe and America. Before all of them he read papers at one time or another and in many cases he served in some capacity on their governing bodies. He became a Fellow of the Royal Society in 1912, a Fellow of St John's College, Cambridge in 1898, and of Corpus Christi College, Oxford in 1912, while in 1919 he received the Honorary D.Sc. of Manchester, his own first University.

In spite of his great ability, his productiveness and energy, and his undoubted achievements, in spite even of what he himself has called his 'arrogance', McDougall was not free from disappointments and from frequently recurring doubts as to the value of his own work. Compelled, probably by his nature, to play for the most part the role of a lone fighter, he founded no school, as so many other modern psychologists of comparable eminence have done, though this fact is probably also due in part to an adverse intellectual atmosphere (at Oxford and at Harvard) and to the circumstance that his psychology was not especially associated with any particular method or technique of investigation. He sometimes felt very keenly the absence of any recognized band of followers who accepted the main principles he had laid down. He was even at times (as some thought) unduly sensitive of what he considered to be the neglect of his teaching on the part of colleagues and students and their failure to make adequate use of it in their own contributions. He perhaps failed to realize the full degree of the influence he was exerting, overestimated the differences between himself and others, and saw stark incompatibilities in certain cases where further reflexion might have revealed possibilities of bridging the gap between his views and those of fellow-workers. But such doubts and misconceptions, in so far as they existed, were rarely if ever allowed to interfere, either with his personal relationships, which, if seldom intimate, were almost invariably friendly, or with steady, persistent work at the tasks which he had set himself. The disappointments, the difficulties, the doubts, such as they were, were for himself alone; his achievements were for the benefit of his chosen science, and through that for mankind at large. For he was acutely conscious of the important role that psychology ought to play in human affairs, and at the same time regretfully aware that it was at present unfitted to play this role, partly because of its humble and backward position in the hierarchy of sciences, and partly because of dissensions among psychologists themselves. McDougall's life exemplified his teaching: it was a life of single-minded purpose, which guided him throughout the innumerable details of each individual undertaking. Early in his career he determined to devote himself to pure research into the secrets of human nature: this course he consistently pursued until the end. Few men could say with the same full sincerity as he: "In the main I have lived the sort of life which in my youth I judged to be the most desirable: and that perhaps is all a man can properly demand."[1]

J. C. FLUGEL

---

[1] In preparing the above account, I have enjoyed the great advantage of consultation with Prof. Cyril Burt and Dr May Smith, to both of whom I am much indebted for kind help and comment.

THE CO-OPERATION OF SOCIAL GROUPS
PART II

By D. M. CARMICHAEL

From the Cambridge Psychological Laboratory

- IV. *Further experiments in constructive social thinking* (pp. 329–331).
  - (i) *Material of Case II: "Co-operation in agriculture"* (pp. 330–331).
  - (ii) *Method of presentation* (p. 331).
- V. *Analysis of results of Case II* (pp. 332–334).
  - (1) *Numerical analysis* (p. 332).
  - (2) *Qualitative analysis* (pp. 333–334).
    - (i) *Permanent co-operation* (pp. 333–334).
    - (ii) *Breakdown* (p. 334).
- VI. *Case III* (pp. 335–336).
  - (i) *Material of Case III: "An anachronism removed"* (pp. 335–336).
  - (ii) *Method of presentation* (p. 336).
- VII. *Analysis of results of Case III* (pp. 336–339).
  - (1) *Numerical analysis* (pp. 336–337).
  - (2) *Qualitative analysis* (pp. 337–339).
    - (i) *Increasing co-operation* (pp. 337–338).
    - (ii) *No real co-operation* (p. 338).
    - (iii) *'Doubtfuls'* (pp. 338–339).
- VIII. *Summary of results of Cases I, II and III* (pp. 339–340).
- IX. *Discussion* (pp. 340–344).
  - (1) *Material* (pp. 340–342).
    - (i) *The questionnaire* (pp. 340–341).
    - (ii) *Detailed and generalized descriptions* (pp. 341–342).
  - (2) *Some possible criticisms* (pp. 342–344).
    - (i) *Concreteness* (pp. 342–343).
    - (ii) *Psychological remoteness* (p. 343).
    - (iii) *Real and described situations* (pp. 343–344).
- X. *Summary* (p. 344).

IV. FURTHER EXPERIMENTS IN CONSTRUCTIVE SOCIAL THINKING

In the case already considered (this *Journal*, xxix, 206–31) conditions which might be held to facilitate or to hinder the co-operation of social groups were presented in a personal and particular form. The two instances which remain to be discussed present evidence for judgement in a more generalized form. The analysis of the results secured can be set out briefly, and it will be seen that interesting differences are found between these cases and the earlier one.

(i) *Material of Case II: "Co-operation in agriculture"*

Yesterday the culminating step was taken in a series of negotiations which, as most people know, have been steadily progressing for several months. Mr Ardern, representing the Union of Agricultural Engineers, and Mr T. Smith, of the Amalgamated Association of Agricultural Labourers, set their signatures to Articles of Agreement which, it is hoped, will bring these powerful, but widely different groups into close and fruitful co-operation. In many quarters it has for long been maintained that the different organized sections of any complicated industry must be brought to unite if the best interests of each section are to be secured. It is undeniable that in the past the increasing invasion of agriculture by the skilled technical worker has produced a great amount of irritating friction. In the mass, the agricultural labourers of this country represent a solid, but excessively conservative element of the community. The alert engineer, on the other hand, marches in the van of an incalculably rapid and bewilderingly varied progress, and is perhaps apt, on that account, to be of a radical frame of mind which welcomes change, and tends to regard attachment to ancient methods and practices as timid and bovine....

A survey of the Articles of Agreement makes it clear that one of the outstanding hopes of those who have worked for this union of groups is that thereby the bargaining power of the labouring class in agriculture towards securing for themselves better economic conditions will be enormously increased. Indeed, a superficial study of the Articles leaves one with the impression that the Association of Agricultural Labourers have far more to gain from the proposed co-operation than have the Union of Agricultural Engineers. But deeper considerations make it obvious that no such impression can be long maintained. Ultimately the interest of all sections of any organized industry must be congruent, and any amelioration of the conditions of any one section must work for the betterment of all. This fact, however, is frequently obscured, and the achievement of this desirable end is hindered, by sectional jealousy and the biassed demands of short-sighted agitators. It is a welcome sign of the times that the Agricultural Engineers clearly recognize that their own good fortune must be founded upon the bed-rock principles of justice and cannot be finally and firmly secured save by active concern for the lot of their less fortunate brothers in industry....

Already critics are not lacking who endeavour to point out that these Articles contain the seeds of division. They say, with undoubted truth, that the mode of organization of the Engineers differs greatly from that of the Labourers. The former, with their vastly greater internal differentiation, have developed along lines which allow far greater scope and power to the expression of local opinion. Their leaders are democratically elected and are, in a true sense as regards their policy, the servants of those whom they lead. The latter, varying far less among themselves, and from district to district, have never developed local organization to a high degree. The Labourer leaders have come to the front mainly on a basis of their particular personal qualities, have framed their own policy and maintained it, when they have been successful, primarily because they possess the power to persuade or to command.

> It is maintained—perhaps we should say rather it is hoped—that co-operation between groups whose organization is on the one hand democratic and on the other despotic can never succeed.
>
> Opponents of the Agreement have even sought the aid of 'science'. They have called in the ubiquitous psychologist. Apparently some eminent and ingenious person, armed with methods which it is not easy to understand, has 'demonstrated' that the 'average I.Q.' level in the Union of Agricultural Engineers is 160, whereas the corresponding measure in the case of the Amalgamated Association of Agricultural Labourers is 90. From these mystic figures dire results are prophesied.
>
> The essential rightness of the Agreement remains. We are convinced that we speak for the solid mass of public opinion in this country when we assert that its principles will stand firm against all selfish and narrow-minded attack, an object-lesson of the way in which industrial organization should and must move, not only in this country, but throughout the world.

This passage was written by Bartlett in the form of a hypothetical leading article, and was presented as such to all subjects. In view of the interesting results which it had already yielded,[1] we decided to test it more extensively.

It differs in three obvious ways from Case I: (1) no specific setting is given; (2) instead of a narrative of events, with little explanation, we have a deliberate and generalized statement of possible leading conditions; (3) the groups described are not directly in contact but attempt to secure co-operation through representatives.

(ii) *Method of presentation*

This material was presented to two groups, both academic—one a mixed class of psychology and anthropology students in Cambridge, and the other a class of psychology students in Edinburgh. The average age distribution in both groups was closely similar.

Bartlett's method of presenting the material was followed, so as to justify comparison of results. The passage was read once to each group as a whole, and any subject who wished to do so was allowed to refer to the text as much as he pleased in giving his answer.

The instructions which followed the reading were these:

> Assuming the facts are as reported, and ignoring any *logical* excellencies or failings in the argument, try to base an opinion, based as far as possible on *psychological grounds* only, as to whether the co-operation sought between these two groups is likely to be successfully maintained, or must inevitably break down. State the basis of your opinion in psychological terms, but without argument.

---

[1] "The co-operation of social groups" (1938), *Occupational Psychol.* XII, 30–42.

V. Analysis of results of Case II

(1) *Numerical analysis*

Table XIV

Group	Permanence	Breakdown	Conditional	Total in group
I. English Academic	15	30	—	45
II. Scottish Academic	40	70	5	115
Total	55	100	5	160

Table XIV gives a simple numerical analysis of the results of Case II, and Table XV the same results expressed as percentages.

Table XV

Group	Permanence %	Breakdown %	Doubtful %
I. English Academic	33·3	66·7	—
II. Scottish Academic	34·8	60·8	4·4
Both groups taken together	34·4	62·3	3·3

The similarity in distribution in these two groups is very striking, and is almost identical with that which Bartlett obtained using the same material. From two English groups, one academic and the other a W.E.A. group (consisting of persons without University training, but with a wider experience of practical affairs), Bartlett got an overall distribution of 35 % for permanence, 61 % for failure, and 4 % doubtful. There was, however, an interesting difference between his two sets of subjects. In the University group opinion was exactly equally divided between probable breakdown and probable permanence. In the practical group the figures were—for permanence 20 %; for breakdown 80 %. No member of the practical group adopted a neutral attitude. Whether a further extension of the survey so as to include a group corresponding to this practical group of Bartlett's would significantly alter the overall distribution of opinion has still to be determined.

In the Scottish groups there is one small internal difference which seems worth noting: of the 70 students who predicted failure, 40 were men and 30 women; and of the 40 who predicted permanence, 17 were men and 23 women—i.e. the women were more inclined than the men to predict permanence, and less inclined to predict breakdown.

(2) *Qualitative analysis*

(i) *Permanent co-operation.* In Table XVI are classified the reasons given for predicting *permanent* co-operation. Table XVII shows the relative emphasis on each reason calculated as described above (p. 219).

Table XVI

Reason	E.A. (15)	S.A. (40)	Both groups (55)
I. Overlap of economic interests	12	26	38
II. Intelligence of engineers	3	9	12
III. Workers of the world uniting	—	6	6
IV. Education reducing difference in I.Q.	—	4	4
V. Effective contact only between leaders (difference in I.Q. small)	—	3	3
VI. Difference in I.Q. irrelevant	—	8	8
VII. Agreement once reached unlikely to break down	—	3	3
VIII. Other 'isolated' reasons	9	3	12
Total	24	62	86

Table XVII

Reason	E.A. %	S.A. %	Both groups %
I	50	41	44
II	12	15	14
III	—	10	7
IV	—	6	5
V	—	5	4
VI	—	13	9
VII	—	5	4
VIII	38	5	14

As in Bartlett's results (when 55% of the subjects who predicted permanence did so on the ground that an overlap of economic interests would defeat all tendency to disruption), the economic factor is by far the most heavily weighted. The superior intelligence of the engineers is considered likely to put them in a dominant position, but no mention is made, in any of the answers analysed above, of a "swing towards democracy among the labour class", such as Bartlett's subjects predicted. Instead, the engineers are expected to adopt a rather paternal attitude, continuing to concede all advantages to the labourers, and showing patience and adaptability.

In Factor VII we see employed the generalization examined in Case I above.

The 'negative reason' that difference in I.Q. is irrelevant, takes two forms: (*a*) intelligence tests are unreliable, and (*b*) differences in intelligence are not usually an obstacle to co-operation.

The 'isolated' reasons included under VIII are reasons which occur only once and are thus unclassifiable.

(ii) *Breakdown.* In Table XVIII are given the reasons on the basis of which breakdown is predicted, and in Table XIX the relative emphasis on each reason.

Table XVIII

Reason	E.A. (30)	S.A. (70)	Total (100)
I. Difference in intelligence	18	35	53
II. Differences in organization and methods of representation	22	16	38
III. Difference in way of life and general outlook	12	60	72
IV. Conflict of interests	8	26	34
V. Other 'isolated' reasons	—	3	3
Total	60	140	200

Table XIX

Reason	E.A. %	S.A. %	Both groups %
I	30	25	27
II	37	12	19
III	20	43	36
IV	13	19	17
V	—	2	2

Differences in organization are ranked as most important by the English group, and as least significant by the Scottish; differences in way of life, outlook and temperament are so strongly emphasized by the Scottish group as to appear most important when both groups are taken together.

Of those of Bartlett's subjects who predicted breakdown, 33% did so because of the difference in I.Q.; 26% held that breakdown must occur because one group is democratically led and the other not; 20% stressed the conservatism of the Labourers, and 15% predicted a clash of interests owing to the increasing use of machinery. But to the practical group the difference in I.Q. appeared far more important and far more antagonistic to co-operation than to the academic group.

In spite of general agreement as to the probability of breakdown, each group shows a tendency to select some particular set of factors which will ultimately bring it about. This strongly suggests that the selection of specific conditions as favourable or unfavourable to social co-operation is itself specially determined. Further experiment, with different groups, will have to be carried out both with these generalized conditions and with others.

VI. CASE III

(i) *Material of Case III: "An anachronism removed"*

The Amateur Rowing Association and the Henley stewards are to be congratulated on their decision to remove the invidious bar which they have hitherto maintained against manual workers. Most oarsmen have long regarded as a blot on the sport the section of the amateur definition which excluded every one "who is or has been, by trade or employment for wages, a mechanic, artisan, or labourer or engaged in any menial duty". So long ago as 1890 a rival body, the National Amateur Rowing Association, was formed on the initiative of such indisputable amateurs as Dr F. J. Furnival, the Reverend Sidney Swann, and Prebendary Propert, with the object of providing opportunities for rowing for amateurs without distinction of occupation. Encouraged by rowing men such as Lord Desborough, Lord Iveagh, and Captain R. C. Bourne, the N.A.R.A., which was reconstituted in 1930, has grown until its affiliated clubs now number about a hundred with some thousands of active oarsmen. None of its crews, however, are eligible to compete at Henley or at any of the A.R.A. regattas, and none of the members of the A.R.A. are allowed by their association to row in regattas organized by the N.A.R.A., although, except for the class bar, the amateur definition of the two associations is in substance the same. This anachronism, reminiscent of obsolete social distinctions, will cease to exist at the end of the present season, when all genuine amateurs, whatever their position in life, will be admitted to compete in all amateur regattas, including Henley, whether held under the auspices of the A.R.A. or of the N.A.R.A.

There can be no doubt that the sport has suffered from the exclusion of a great source of promising material from all possibility of ever taking part in the classic events. The rigid division into sheep and goats who must not be allowed to mix, inevitably interfered with the development of rowing and with the promotion of good fellowship, one of the chief objects of every sport. Cases such as that lamented in a recent letter from the President of the Southampton University Boat Club have not been rare. Rowing men whose memories are long enough will remember the unpleasant effect produced when an N.A.R.A. eight was refused permission to compete for the Cup offered by the King at the regatta held to celebrate the end of the War. In the event this Cup was won by an Australian crew, to whose members the bar against manual workers must have seemed incomprehensible, since no such class distinction would be tolerated in any sport in Australia or any of the other Dominions. One of the worst effects of the rule now rescinded is that it has deterred Dominion crews who have wished to measure themselves against English oarsmen from competing at Henley. When they have been told that their entry could not be accepted because some of them worked for their living with their hands they have been at first incredulous and then contemptuous of what they regarded as another exhibition of English snobbery. In this country, where its origin has been understood, as well as the natural desire to preserve the family party atmosphere characteristic of English regattas, the bar has caused very little ill-feeling. The N.A.R.A. would be the first to acknowledge the sympathy and the help they have received both from the A.R.A. and from the Henley stewards. Now that an outworn distinction has been swept away, there is every prospect of increasing co-operation between the two associations, to the benefit of both and of English rowing as a whole.

This passage appeared as a leading article in *The Times* in June 1937, and is reproduced here by kind permission of the Editor. It was chosen as experimental material primarily in order to check the predictions made on the basis of it against the actual course of events. In general style (being a judicious mixture of historical facts and conventional generalizations) it is eminently suitable for presentation to groups both of experts and of non-experts and for effecting a comparison between them.

(ii) *Method of presentation*

The material was presented to three groups—an academic non-rowing group, a non-academic non-rowing group, and an academic rowing group. The first two groups included both English and Scottish subjects, whose opinions are so similar as not to require separation. The third is a group of Cambridge rowing men. Some variation in the methods of presentation proved necessary. Results were obtained from the first and third groups partly by individual and partly by group presentation, and from the second only by correspondence. To what effect this affected the answers is not known. As before, each subject was allowed to refer to the material as much as he liked in dealing with it. The question asked was:

> Do you think that there is every prospect of increasing co-operation between the two associations? State your reasons.

VII. ANALYSIS OF RESULTS OF CASE III

(1) *Numerical analysis*

Table XX

Group	Yes	No	Doubtful	Total
I. Non-rowing academic	80	22	18	120
II. Non-rowing non-academic	16	5	9	30
III. Rowing academic	10	2	18	30
Total	106	29	45	180

Table XX gives a simple numerical analysis of the results of Case III, and Table XXI the same results expressed as percentages of the total number of responses.

Table XXI

Group	Yes %	No %	Doubtful %
I	66·7	18·3	15
II	53·3	16·7	30
III	33·3	6·6	60

The non-academic non-rowing group is one of upper working and middle-class persons, and the numerical analysis reveals the variation of opinion between them and the first group. It is, however, rather misleading as regards the third. The first two groups, without technical knowledge, are inclined to accept the phrase 'increasing co-operation' without much analysis, and to give a straightforward affirmative or negative answer. The doubtfuls in each of these groups are subjects who are unwilling to be definite because their knowledge of the facts is extremely limited. In the expert group, however, while the number of direct negatives is remarkably small, the number of doubtfuls is as high as 60%. This is due not to lack of knowledge of relevant data, but to excess of it. The doubting experts give an extremely detailed account of what is likely to happen, and owing to the complexity of the situation as they see it, their answers are strictly neither affirmative nor negative.

(2) *Qualitative analysis*

(i) *Increasing co-operation.* Table XXII shows the causes considered likely to bring about increasing co-operation between the two associations.

Table XXII

Cause	N.R.A. (80)	N.R.N.A. (16)	R.A. (10)	Total (106)
I. Common interest in sport eliminates class-distinctions	39	10	8	57
II. Co-operation to benefit of both associations, of this sport as a whole	19	4	2	25
III. Gradual disappearance of class-distinctions	24	5	7	36
IV. Personal contacts	10	9	2	21
V. Other causes	8	2	6	16
Total	100	30	25	155

Increasing co-operation is predicted in all groups mainly on the basis of the conventional generalization that "sport eliminates class distinctions". The practical (non-rowing) group backs this up closely by expressing confidence in another generalization—that "men individually are not snobs, though in a group they may act as such". What is evidently meant is that when two groups which have previously stood apart from one another are brought together, the personal contacts of individual members of each group will go a long way towards breaking down class distinction, and creating a real understanding.

The disappearance of class barriers throughout this country as a whole is most clearly appreciated as relevant to this case by the two academic groups.

The benefits of co-operation are envisaged in terms of prestige rather than of economic gain, though the latter receives some attention.

(ii) *No real co-operation.* Table XXIII shows the frequency with which each group selects different conditions as being likely to prevent any real co-operation.

Table XXIII

	Causes likely to prevent any real co-operation	N.R.A. (22)	N.R.N.A. (5)	R.A. (2)	Total (29)
I.	Sheer snobbery	2	—	2	4
II.	Incapacity to see each other's point of view	4	—	—	4
III.	Rivalry within nominally united body	15	4	—	19
IV.	"The peculiar rowing tradition"	2	—	—	2
V.	A.R.A. removed bar only in deference to public opinion	2	3	—	5
VI.	Nature of the sport of rowing itself (no real analogy with other sports)	—	—	2	2
	Total	25	7	4	36

Sheer snobbery on the part of the A.R.A., in the sense that they will refuse to recognize the N.A.R.A. is little anticipated. On the other hand, condescending recognition on the one hand, with self-assertiveness and an inferiority complex on the other, are considered to be the most probable causes of rivalry and jealousy, and of a merely nominal union.

Two rowing men mention that, "although sport is nominally a social link because of the close contact into which sportsmen come, this is not so in rowing. All intercourse is merely social." In other words, the crews train and race as separate, self-contained units, and only meet under the normal conditions of social life. Both the men who hold this opinion also state that "in the elegant atmosphere of Henley the N.A.R.A.'s would be marked men, and their presence would be disagreeable to undergraduates who consider rowing a University monopoly, and try to separate themselves from the working class".

(iii) '*Doubtfuls.*' The rowing group are alone of interest here. A typical answer would run somewhat as follows: "In England there are two distinct standards of A.R.A. rowing—(*a*) Henley and London, (*b*) Northern provincial.

"(*a*) *Henley*—little co-operation likely, either in the sport itself, or socially. Few N.A.R.A. crews (for reasons of time and expense) can hope to compete successfully with University and Tideway (University trained) crews, and so few are likely to row in the regatta. Nor will there be much social co-operation, for even now there is little contact between the various clubs, and one meets only people one knows already, and their friends.

"*Other Thames regattas*—the co-operation depends on the number and rowing strength of the various N.A.R.A. clubs on the Thames.

"(*b*) *Northern regattas*. The A.R.A. and N.A.R.A. clubs should join together very well. All are of the same social class, financial position and ability as oarsmen; and since the same crews compete against each other on several occasions at different regattas, they cannot help getting to know one another, and increasing social intercourse seems highly probable.

"The standard of N.A.R.A. rowing might be improved if A.R.A. men would row for N.A.R.A. clubs, but it is unlikely that they will do so, even though the bar has been lifted."

This eclectic answer is interesting as showing (*a*) that the experts are far less guided by the governing generalizations of the leading article, than are the non-experts, (*b*) that they tend to be much more specific in their predictions, and (*c*) that they introduce a factor which is of primary importance in securing co-operation, and which is omitted from the leading article, and for that reason by all the non-experts, viz. the improvement of the standard of rowing in N.A.R.A. clubs.

VIII. SUMMARY OF RESULTS OF CASES I, II AND III

The analysis of Case I reveals how in dealing with certain group relationships, five selected groups of subjects made either implicit or explicit use of the following main general principles:

(1) (*a*) Groups are seldom, if ever, capable of disinterested conduct, and the co-operation of groups depends on how far the interest of the group is furthered thereby, and the pride of the members in their own group, which flatters their individual self-esteem, satisfied.

(*b*) The causes which facilitate and consolidate the union of religious groups (and probably of most other groups as well) are generally very far removed in tone from so-called Christian motives.

(2) (*a*) A group surcharged with emotion, or one which has just passed through a period of emotional stress, is generally more willing to co-operate than when in its normal state.

(*b*) The co-operation thus secured is apt to break down unless stabilized by other causes.

(3) Most groups display a natural conservative tendency to unite against disruptive factors.

(4) Any two groups which fulfil much the same function in a community and which, while vigorous, resist all attempts at union, will co-operate immediately they are so reduced in numbers or vitality as to be faced with destruction.

(5) The central control which a large organization exercises over its sub-groups facilitates the co-operation of these groups.

(6) Religious differences, being the outcome of inescapable temperamental characteristics, cannot be resolved by any externally imposed union.

(7) (*a*) Country and seafaring folk derive from their way of life a temperament and cast of mind wholly unsuited to consort with an urban population.

(*b*) This difference is more than one of economic organization, but the first step towards securing co-operation between an urban group and a group of country or fisher folk would be a change in economic organization.

(8) *Either* A rebel 'within a tradition' is always supported, but becomes ineffective once he sets himself up as wholly independent.

Or A 'cause' receives support only in so far as, and as long as, it makes an eloquent appeal to a man's natural egotism and flatters his belief in his own importance.

While Case II does not yield any generalizations over and above those selected as leading conditions by the experimenter, it reveals very effectively the importance which is placed upon those generalizations by different social groups. Further experimentation is required to show how the use of those generalizations is itself specially conditioned (p. 334 above). Case II shows also that, while there may be close agreement as to the chances of co-operation or breakdown in any given situation, there may be considerable disagreement as to the precise psychological conditions which bring about co-operation or breakdown.

Case III is interesting chiefly as illustrating how 'psychological remoteness' determines the manner in which subjects deal with a given situation (see below, p. 343).

The accuracy of the predictions made by the two groups of non-expert subjects, and by the group of experts, will in due course be checked against the actual course of events.

IX. Discussion

This paper has been written primarily to illustrate a method. I shall therefore conclude with some general remarks on the method, with special reference to our wider scheme of research, and with a view to meeting possible criticisms.

(1) *Material*

(i) *The questionnaire.* The method has, for our purpose at least, two advantages over most questionnaires: (1) By referring to a particular concrete case, the experimenter can be reasonably sure that his questions have the same essential meaning for his subjects as they have for himself. (2) If we ask an individual a set of isolated questions concerning the nature and relative importance of the factors governing group relationships, then we can scarcely hope to know what aspects of his experience he has recourse to in framing his answers. If, however, we describe a situation in which certain postulated factors are *possibly* relevant, then we can observe what selection and weighting of these factors occurs, and what further factors, if any, are imported. By thus controlling the types of situation used as frames of reference, and comparing the response of different social groups to different situations, we may perhaps discover how this all-important selection and weighting of governing conditions is itself specially determined (cf. p. 334 above).

(ii) *Detailed and generalized descriptions.* The experiments just described, however, only justify us in asking what important differences are found between those cases in which the conditions of possible co-operation have to be inferred from a considerable amount of particular detail (Case I, p. 210) and those in which they are presented in a generalized form (Case II, p. 330 and Case III, p. 335). The differences are as follows:

(1) A quick return can best be obtained by the use of examples similar to Case II. The issue being stated in terms of a few generalizations, most subjects solve it simply by utilizing these, and their opinions can be quickly classified. By the further use of similar material it should be possible speedily to determine the frequency and distribution of *any* conventionalized generalization throughout a given social group.

In so far as widespread opinion (either *true* or *false*) concerning the factors governing social co-operation affects the chances of co-operation, an enquiry into the frequency and distribution of such opinion will have practical value.

(2) Very detailed descriptions such as those of Case I, on the other hand, yield results which require more careful sifting. This is chiefly because, as noted above (p. 220), the psychological principles on which many solutions are based are frequently not explicitly stated, but have to be inferred and formulated by the experimenter.

Case I (unlike Cases II and III) yielded in addition to these unformulated principles a number of unsolicited generalizations.

Very detailed descriptions thus seem to have a double value. Firstly, they reveal the presence of conventionalized generalizations the frequency and distribution of which can be more fully determined in the manner described above—i.e. by embodying them in less detailed descriptions, and presenting them to selected groups of subjects. Secondly, when subjects are asked to state the probable upshot of a situation described in great detail, some of them, avoiding easy generalizations, base their opinion largely on specific conditions selected from the given context.

Underlying the specific conditions so selected it may sometimes (as in Case I above, pp. 220, 221 and 222f.) be possible to discover a principle of fundamental importance. This principle may never have been consciously formulated as a principle by any of the subjects in question, yet in so far as they make implicit or unconscious use of it, we can infer that they have had, in real life, empirical evidence of its working and are able to recognize factors relevant to it. A more extended use of material similar to that of Case I should therefore enable us profitably to tap the actual working experience of individuals within selected social groups and so to gain some indication of the *real* factors which govern social harmony and discord.

To sum up: (1) generalized descriptions are most suitable for showing how people think about social questions (p. 340); (2) very detailed descriptions are most likely to provide us with the guiding principles essential for profitable observations and experimentation in the field (p. 208, ii).

(2) *Some possible criticisms*

(i) *Concreteness.* It will probably be urged that in spite of its claim to concreteness our material must, by its very nature, remain viciously abstract. All descriptions being essentially the result of selection and interpretation, no description can ever represent an actual situation in its full complexity. In asking our subjects to deal with described situations, we are therefore asking them to deal with biased and incomplete pictures of the actual.

The reply to this is twofold: (1) In dealing with an actual situation in which a practical decision has to be made, every person has to exercise some kind of selection. He cannot hope to know *every* aspect of the situation, and from those of which he is aware, he has to pick out the relevant from the irrelevant. In other words, the issue is decided, not in terms of the whole situation but in terms of a few selected factors. In the same way, when we are contemplating others acting in a real situation, we must, if we wish to predict the probable course of events, select those factors which we consider will govern the agents' actions. (2) In confronting our subjects with a described situation, and asking them to predict how they themselves or others would act under those conditions, we are merely presenting them with a case in which some of this essential work of selection has already been carried out. As we have seen (p. 340), it is, for the purposes of our enquiry, a definite advantage to be able thus to keep a check on our subjects' summing-up of the situation, and to observe any further selection or importation of data which may determine their predictions. The risk of suggesting the probable issue has simply to be taken.[1]

(ii) *Psychological remoteness.* It may be objected, again, that there is little profit in collecting opinions from subjects concerning situations from which they are 'psychologically remote'. In this context, a subject would be psychologically remote from any situation which describes group-relationships of which he has had little or no experience either as an observer or as a participant.

The experiments already carried out (most notably Case III, above) indicate that most individuals show remarkable confidence in dealing with situations which are psychologically remote in this sense. As we saw, in predicting the probable course of events, they readily fall back on conventionalized generalizations concerning group relationships. The practical value of determining the nature, frequency and distribution of these generalizations may be disputed, but an enquiry into the extent to which such generalizations govern the way in which people think about social questions would be of undoubted theoretical interest. A beginning at one end of the scale has already been made by collecting the opinion of a number of subjects concerning a group relationship in which they are actual participants. The investigation will be continued with material of ever-increasing remoteness, and a report will be published in due course.

(iii) *Real and described situations.* Finally, it may be maintained that *every* described situation, just because it is described, must remain psychologically remote. No matter how closely a described situation may concern us, and no matter how vivid our description, the description can never possess the force and urgency of the actual; and therefore our reaction to the description can give no sure indication of our reaction to the real situation. Implying as it does, that we have neither knowledge nor control of the key factors which govern our behaviour as members of a group, this criticism raises issues which it is impossible to discuss here.

---

[1] Cf. H. Cantril, "The prediction of social events", *J. Abn. Soc. Psychol.* (1938), xxxvi: "The uniformity of predictive judgements regarding a social event is in part, but not wholly, proportional to the degree to which the specific factors leading up to an event can be resolved into a consistent and meaningful pattern." Cantril reaches this conclusion after using the *questionnaire* method to obtain, from selected individuals belonging to eleven different social groups, their opinion as to the probable upshot of a variety of contemporary situations familiar to an informed American. Just how far a deliberate resolving of these situations by the experimenter into "consistent and meaningful patterns" biases subjects' predictions could be at least partly determined by embodying Cantril's (or other) questions in descriptions such as Case III above, and using them with similar groups of subjects.

For the present we can only point out that the criticism, if pushed to extremes, would apply to *all* investigation in psychology which uses situations built up under experimental conditions; and, once again, that this particular investigation is a preliminary one, expressly devised to give direction to, and be supplemented by, other methods.

X. SUMMARY

I. Three interdependent lines of research are proposed:

- (*a*) Direct observation of limited communities.
- (*b*) Controlled experiments in the field.
- (*c*) Experiments in constructive social thinking.

The conclusions which Bartlett reached after a series of experiments of the last type are summarized.

II. Three further experiments in constructive social thinking are described, and their results analysed.

III. An attempt is made to assess the significance of these results both in themselves and in relation to the wider scheme of research. The discussion therefore is concerned mainly with the theoretical validity and the practical value of the method of investigation employed.

I have to thank Prof. F. C. Bartlett for his inspiration and help; and Prof. James Drever for his kindness in allowing me to conduct group experiments in his laboratory in Edinburgh.

(*Manuscript received* 15 *April* 1938)

VARIABILITY OF JUDGEMENT AND 'STEADINESS' OF CHARACTER

By JEAN D. CUMMINGS

I. *Introduction* (pp. 346–347).

II. *The experiment* (pp. 347–350).
  (1) *The block-in scale* (A) (p. 347).
  (2) *The hypothetical situation scale* (B) (pp. 347–348).
  (3) *The Marston scale of introversion-extraversion* (p. 348).
  (4) *The tests* (pp. 348–350).
  (5) *Subjects' introspections* (p. 350).

III. *The test results* (pp. 350–359).
  (1) *Reliability* (pp. 350–352).
  (2) *Comparison of continuous and day-to-day work* (pp. 352–353).
  (3) *Intercorrelations of error, speed and variability* (pp. 353–355).
  (4) *Intercorrelation of the tests* (pp. 355–359).

IV. *Rating scale results* (pp. 359–363).
  (1) *The block-in scale* (A) (pp. 359–361).
  (2) *The hypothetical situation scale* (B) (pp. 361–362).
  (3) *The Marston introversion-extraversion scale* (p. 362).
  (4) *Reliability of the scales* (p. 362).
  (5) *Intercorrelation of rating scale results* (pp. 362–363).

V. *Correlation of rating scales and test results* (pp. 363–367).
  (1) *Ratings and variability scores* (pp. 363–366).
  (2) *Ratings and error scores* (p. 366).
  (3) *Ratings and speed* (p. 366).
  (4) *Correlation of variability with the whole block-in scale* (pp. 366–367).

VI. *Summary of main results* (pp. 367–368).

References (pp. 368–369).

I. INTRODUCTION

OF the problems of vocational selection, one of the most difficult seems to be that of forecasting the 'reliability' or 'conscientiousness' of an individual and his habits of work. The measurement of his intelligence and special abilities gives no measure of his tendency to work conscientiously and thoroughly at a monotonous task, or at a task calling for concentrated attention to detail.

In the present experiment it was thought that the '*w*' factor of Webb's(1) research, which he described as being in close relation to 'persistence of motives', might be revealed in the various features of the work curve for a group of subjects. In particular, it was thought that its relation to the following measures might be of interest:

(1) The average error of judgement in each 3 min. period for 20 periods.

(2) The average variation of judgement in each 3 min. period for 20 periods.

(3) The average speed of judgement in each 3 min. period for 20 periods.

With regard to (2) above, Cox(2) has shown that in tests of manual skill, the daily variability of his more able subjects was less than that of his less able subjects. Walton(3), in short-period work with children, found that small amplitude of oscillation in two tests of motor abilities and two of cognitive processes correlated with estimates of 'steadiness' of character. There is therefore ground for enquiring whether variability (in this case variability of judgement) may not be correlated with low 'steadiness', i.e. correlated negatively with '*w*'.

The analysis of oscillation has shown '*o*' to be a general and not a specific factor, but it cannot be assumed that variability in more complex tests is anything more than a characteristic specific to each test, caused by skill, interest and other factors. Hollingworth(4) could produce no conclusive evidence for a variability factor in such tasks as tapping, colour naming and in the variations of the pulse rate, although Asch(5), repeating the experiment, but eliminating the pulse-rate scores, found a systematic though slight factor common to the tests (average r was 0·1972).

With regard to measures of 'persistence', tests so far devised seem to have shown that an individual's persistence score is specific to the task he does. May & Hartshorne(6) have shown this conclusively in short-period tests. But there is a great deal of difference in the usage of the term ‘persistence’. May and Hartshorne use it to mean “the ability to resist the tendency to go more and more slowly”. Crutcher(7) uses it to mean the tendency to continue definite activities in the face of difficulties. Howells(8) measures it by the amount of endurance with which a person can bear increasing discomfort and fatigue. One of the chief difficulties seems to be that in miniature situations the situation is not alike—does not mean the same—for every subject.

One aim of this experiment was to see, by independent estimates, whether several uses of the word ‘persistence’ do in fact denote a group factor—whether for instance ‘persistence in routine’, ‘persistence of interests’, ‘persistence of principles of conduct’, etc., are highly intercorrelated: finally, given any reliable estimate of persistence, to see whether this was revealed in any feature of the work curve.

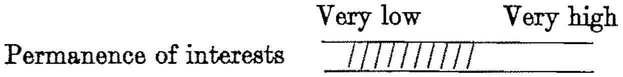
II. THE EXPERIMENT

The subjects were eighteen women undergraduates in their second year at college. Each subject was asked to obtain three judges from among intimate friends, each of whom would assess her in the rating scales; and to assess herself on the same scales independently.

The procedure of testing has been previously described in more detail(9), but was, briefly, as follows:

(1) *The block-in scale* (A)

In this scale the rater was asked to assess the subject in thirty traits, selected as ‘*w*’ traits, plus additional traits relating more closely to work qualities and persistence in general. The traits were listed on the left-hand side of the page and on the right was a series of blanks, which the rater could block in to any degree she thought correct. Thus:



No attempt was made to control the distribution of the values because only a rank order was desired.

(2) *The hypothetical situation scale* (B)

In this scale the rater was given the following instructions:

“Imagine the subject is placed in the following situations and put a tick opposite the action or actions you think she would be most likely to take. If you are unsure, or if you have anything to add, put it in the space left for REMARKS.”

The five questions were designed to measure by particular examples:

Persistence in routine work (checking figures).
Persistence of principles (steadfast support of a society).
Persistence of interests (length of time hobbies last).
Persistence of intentions (changing holiday plans).
Persistence of moods (persistently cheerful or despondent).

The questions correspond to the five persistence traits in scale A, but in this scale B it was thought desirable to describe the traits in specific situations with which anyone might be faced.

(3) *The Marston scale of introversion-extraversion*

A measure of introversion was obtained by using the Marston scale, which although criticized by Guilford & Hunt(10) as being based on the emotional aspect of the introversion-extraversion distinction, was thought to be sufficiently satisfactory to obtain a rank order in a research upon only a small number of subjects.

(4) *The tests*

Testing was individual and was carried on for 5 days a week for 3 weeks, Saturdays and Sundays being omitted. On the first day there was a 3 min. practice period in each of the four tests. On each of the other 4 days of the first week the subjects did continuous work for half an hour, doing one test on each day. In the remaining 2 weeks they did 3 min. on each test each day, making a 12 min. testing time each day. The total testing time for each subject was 4 hr., neglecting of course the time taken between the tests. The tests were:

The arm-movement test (I). The material for this test consisted of two boards, each with two parallel grooves along which small square blocks could be moved and the distance registered in millimetres. Clamps could be inserted at any point in the scale to limit the distance the blocks could be moved. Two such boards were placed before the subject, so that there were four parallel grooves. The subject, starting with the topmost groove, held two small blocks on either side of a central clamp, and with her left hand ran a block as far as it would go away from the clamp. She then ran the right-hand block in the opposite direction to what she considered was an equal distance. She made the judgement with closed eyes, opening them between judgements to see the amount of error she had made. She ran the blocks along the lower grooves in the same way, and then immediately returned to the topmost, and so on without interruption. She was instructed to choose her own speed of work, which would not influence her score in any way. The speed of each judgement was, however, recorded. The length of the standard distance, i.e. the left-hand run, was varied slightly in each case, so that a separate judgement had to be made for each groove.

The ball-and-slot test (II). The material in this case consisted of a series of boxes; the subject placed a ball in the top of the box, allowed it to run along the top of a metal ledge and from there drop into one of twelve slots, marked from 1 to 12. The subject tried in every case to drop the ball into number 6 slot, although the preliminary run along the metal ledge made it more difficult to do this than to drop the ball into numbers 1 to 5. The subject worked, as in test I, without interruption and at what speed she chose. The only restriction was that the box must be held at an angle of 45° or more with the horizontal.

The character judgement test (III). The material used in this test was the same as for the block-in rating scale described above. In the practice given before the test, the subject assessed herself in all the traits and blocked in one complete sheet, so that in the test proper there was no great difference in speed between the first set of judgements and subsequent ones, but only a gradual quickening throughout the test. The subject worked continuously through the sheets, each of which contained the same thirty character traits in which self-assessment was to be made. The traits were, however, arranged in four different orders to prevent the subject remembering the general look of a completed sheet, and in order that judgement had to be made on each trait separately.

Matching greys test (IV). In this test the subject was given a board in which two small squares were cut out. Through the left of these a standard grey was shown, while on the right a revolving disk at the back of the board displayed a series of greys ranging from a grey easily discernible as lighter than the standard to one very much darker. The subject had to judge which one of this series most nearly resembled the standard. Four standards were used in turn, varying slightly in brightness. The subject started from the light and dark ends of the scale alternately, but could reverse the disk as often as she pleased.

From these tests measures were obtained of:

(*a*) The amount of error of each judgement.

(*b*) The time of each judgement.

(*c*) The variability of each 3 min. period. This was measured in tests I and II by the average deviation from the mean of the judgements (neglecting the error). In tests III and IV the variability was obtained by subtracting each judgement from the previous corresponding one and averaging these subtractions regardless of sign for each 3 min. period.

(*d*) The total variability of each subject was obtained for the continuous and subsequent day-to-day work by averaging the variabilities of the 3 min. periods (ten for each type of work).

(5) *Subjects' introspections*

Subjects were asked for introspections in general, and they were also asked:

Which test was the most interesting?
Which test was the most uninteresting?
Which test needed most concentration?
Which test was most physically fatiguing?
In which test was the subject most successful?

III. The test results

(1) *Reliability*

The reliability of the variability scores was calculated in two ways—by correlating the variability scores on odd periods with those obtained on even periods, and by correlating scores on the first five periods with those on the last five periods in both types of work.

The reliability coefficients are shown in Table I, the figures in brackets being the coefficients corrected by the Spearman-Brown prophecy formula to give the reliability of the whole test.

Table I. *Variability*

	Test I		Test II		Test III		Test IV	
	A	B	A	B	A	B	A	B
Odd and even judgements	0·71 (0·83)	0·77 (0·87)	0·76 (0·86)	0·86 (0·93)	0·88 (0·94)	0·81 (0·89)	0·03 (0·06)	0·79 (0·88)
First 5 and last 5 periods	0·78 (0·88)	0·69 (0·82)	0·74 (0·85)	0·83 (0·91)	0·70 (0·82)	0·70 (0·82)	0·14 (0·25)	0·38 (0·55)

It will be seen from the table that while the reliability of tests I, II and III is satisfactory, that of test IV is markedly not so, the reliability of odd and even judgements in the continuous work (A) being as low as 0·06. There are several possible reasons for this low reliability. The subjects from their introspections evidently detected qualitative differences in the grey series, which led them to judge as equal greys which were several shades apart in brightness, but which were nevertheless more 'blue' than the others, and so for this discrepant reason similar. The subject sat always in the same position in the room and in artificial light, but other conditions such as eye fatigue from previous work, or the brightness of light outside the test room, may have altered the day-to-day results. Whatever the reason, it seemed obvious that the test was not sufficiently reliable to justify its use, and it was accordingly discarded.

The other three tests are all of good reliability. All the corrected coefficients are above 0·80, and it seems therefore that both the continuous and day-to-day results are reliable and almost equally so. This finding is a somewhat surprising one, especially in the block-in scale of character judgements, for it is difficult to see why a subject should go through twenty 3 min. periods altering her character judgements and speeding up the rate of making them but in variability should keep so constant a position relative to other subjects that the rank orders at the middle and end of testing should correspond so highly. It would perhaps be easier to understand if there had been a wide range of scores, but the whole length of the block-in scale was only 5 cm., and the range of average scores only 2·57 mm. Why, for instance, should one subject have 3 min. variability scores of 0·185, 0·145, 0·112, 0·172, etc., while another subject scores 0·240, 0·220, 0·216, 0·220, etc.? How can this difference, which is a difference between two series of individual scores, be accounted for?

The reliability of the error scores (there was of course no error scoring in the character judgement test) is shown in Table II. The reliability of test IV (grey matching) is again unsatisfactory, and this test was finally discarded in working out the results. The reliability of the other two tests is satisfactory.

Table II. *Error scores*

	Test I		Test II		Test IV	
	A	B	A	B	A	B
Odd and even judgements	0·89 (0·94)	0·94 (0·97)	0·96 (0·98)	0·95 (0·97)	0·60 (0·75)	0·24 (0·39)
First 5 and last 5 periods	0·63 (0·77)	0·89 (0·94)	0·87 (0·93)	0·89 (0·94)	0·21 (0·35)	0·58 (0·73)

The reliability of the speed scores is shown in Table III. With one exception the coefficients suggest an exceedingly consistent rate of work, consistent both from period to period and between the beginning and end of work. The exception is the continuous work periods of the character judgement test (III). Here the coefficient of 0·68 indicates that although the times of judgement are highly consistent from period to period, there

Table III. *Speed scores*

	Test I		Test II		Test III	
	A	B	A	B	A	B
Odd and even judgements	0·99 (0·99)	0·97 (0·98)	0·94 (0·97)	0·98 (0·99)	0·91 (0·95)	0·97 (0·98)
First 5 and last 5 periods	0 89 (0·94)	0·75 (0·86)	0·83 (0·91)	0·81 (0·89)	0·51 (0·68)	0·79 (0·88)

are individual differences in speeding up or slowing down which affect the rank order when the first and last halves of the test are compared; it is of course not surprising that individuals differed in the way they increased their speed of self-judgement. By the time the day-to-day work was reached individual rhythm of work was evident and remained unaltered except for a slight increase of speed common to all.

It thus appears that all three measures—speed, error and variability—are highly consistent within each of the three activities of the tests. That is to say, a subject tends, relative to other subjects, to keep speed, error and variability consistent throughout one half-hour of continuous work and throughout the following 10 days of daily testing. Differences in improvability and speeding up may alter this relationship somewhat, but in these particular tests their effect is not marked.

(2) *Comparison of continuous and day-to-day work*

Table IV gives the correlations obtained between the continuous half-hour work (A) and the daily 3 min. period work (B), the variability scores

Table IV. *Variability scores*

		Continuous work (A)		
	Test	I	II	III
Day-to-day work (B)	I	0·36		
	II		0·90	
	III			0·80

being considered first. Two of these correlations are significant, showing that in the ball-and-slot and character judgement tests a subject's tendency to vary is very much the same in uninterrupted work and in work spread over 2 weeks. In the arm-movement test (I) there is no such close relationship, and although there are several possible reasons for this they are merely speculative, and the cause cannot be discovered without much more information upon the rate of learning in this test, conditions of boredom, fatigue, etc.

This correlation of variability scores in continuous and in day-to-day work is important chiefly from the point of view of the framing of a diagnostic 'variability' test. If test material can be obtained in which a subject's variability can be assessed as well in an initial long period as in periods spread over a length of time, such material is obviously of more practical use for testing purposes than material which reveals a subject's tendency to variation only after a length of time. From this point of view, the ball-and-slot and character judgement tests are better than the arm-movement test, in which it is quite possible the A type of work is merely an 'absorption' period, i.e. a period into which many extraneous factors enter, to be gradually levelled out and their attenuating influence decreased. The ball-and-slot and character judgement tests are reliable both in long and short period work; a subject's scores on both types of work are highly consistent, and in fact his 'variability' could be diagnosed in one half-hour test with no need for a longer period.

Table V. *Error scores*

		Continuous work (A)	
	Test	I	II
Day-to-day work (B)	I	0·71	
	II		0·93

Table VI. *Speed scores*

		Continuous work (A)		
	Test	I	II	III
Day-to-day work (B)	I	0·63		
	II		0·77	
	III			0·62

The correlations for the error and speed scores are given in Tables V and VI. These correlations are lower, but significant, showing that there is a tendency for an individual's rhythm of work on a test to be consistent whether that work is over a long or short period.

(3) *Intercorrelations of error, speed and variability*

The intercorrelations of error and variability are given in Table VII. It seems evident that the correlation of error scores with variability depends on the nature of the test. In the continuous work of test I highly variable subjects were not necessarily the most accurate. In test II on the other hand the correspondence was from the beginning very high.

Table VII

		Variability			
Test		I		II	
		A	B	A	B
Error I	A	0·27			
	B		0·73		
Error II	A			0·91	
	B				0·94

It is easily understood from the nature of the ball-and-slot test that a subject who was skilful in getting balls into slot 6 would vary less than one who was unsuccessful. A close correspondence of error and variability probably depends therefore upon the nature of the test material, in which success may imply lack of variation (as in the ball-and-slot) or be comparatively independent of it. It must also, evidently, depend upon the degree of knowledge of success given to the subject and the incentive to vary upon knowledge of failure.

The variability and speed correlations were taken between low variability and low speed and were obtained for the first five A periods (A_1), the last five A periods (A_2), the first five B periods (B_1) and the last five B periods (B_2). The results, shown in Table VIII, should therefore show

Table VIII

Variability

Test I (arm-movement)

Speed	A_1	A_2	A	B_1	B_2	B
A_1	−0·15					
A_2		−0·07				
A			*−0·08*			
B_1				−0·41		
B_2					−0·39	
B						*−0·45*

Test II (ball-and-slot)

Speed	A_1	A_2	A	B_1	B_2	B
A_1	0·53					
A_2		0·42				
A			*0·51*			
B_1				−0·10		
B_2					0·22	
B						*0·06*

		Test III (character judgement)					
		A_1	A_2	A	B_1	B_2	B
Speed	A_1	0·05					
	A_2		0·03				
	A			*−0·14*			
	B_1				−0·14		
	B_2					0·48	
	B						*0·06*

how the relationship of variability and speed altered over the 3 weeks' testing time. The italicized figures give the correlations for the whole A and B periods. It seems that the relationship of variability to speed depends upon the nature of the test. In test I there seems to be a slight tendency, more marked in the later testing, for speedy subjects to be the least variable, but this relationship is inverted in the case of the ball-and-slot test; whereas in test III there seems to be no significant relationship.

The error and speed correlations are given in Table IX. Here again, no decisive trend is to be found. Lack of error correlates negatively with slowness in test I, indicating that a fairly high speed is conducive both to lack of error and variability, but this is not the case in test II.

Table IX

		Error											
		Test I						Test II					
		A_1	A_2	A	B_1	B_2	B	A_1	A_2	A	B_1	B_2	B
Speed	A_1	−0·05						0·06					
	A_2		−0·01						0·35				
	A			*−0·24*						*0·25*			
	B_1				−0·44						−0·30		
	B_2					−0·26						0·10	
	B						*−0·37*						*−0·04*

The relations of speed, variability and error seem, therefore, to be dependent on the nature of the test and the progress of it; there is no evidence that any relationship holds between them that is not specific to each test.

(4) *Intercorrelation of the tests*

The question now arises as to whether these measures of variability, error and speed are specific to each test, or are wider factors common to all three tests. The variability results have been correlated at the end of each five periods as before, i.e. at A_1, A_2, B_1 and B_2 as shown in Table X. The coefficients for the whole A and B periods are italicized.

Table X. *Variability scores*

		Test I (arm-movement)						Test II (ball-and-slot)					
		A_1	A_2	A	B_1	B_2	B	A_1	A_2	A	B_1	B_2	B
Test II (ball-and-slot)	A_1	0·32											
	A_2		0·20										
	A			*0·19*									
	B_1				0·60								
	B_2					0·61							
	B						*0·81*						
Test III (character judgement)	A_1	−0·05						0·21					
	A_2		0·12						0·39				
	A			*0·04*						*0·34*			
	B_1				0·20						0·42		
	B_2					0·27						0·25	
	B						*0·37*						*0·47*

The coefficients in this table, although not all significant, are, with one exception, all positive, i.e. correlations taken at four different points in the progress of the test yield positive coefficients throughout. This means that there is a tendency for a subject who is variable at the beginning of test I to be variable also at the beginning of test II and (a little later) in test III, and that throughout the three tests this tendency persists. These correlations, and the fact that on the rankings for the whole B period the average coefficient is 0·55, indicate strongly that in the present sample at least there is either a common factor of variability present; or other extraneous factors, such as overlapping material, are present to account for these coefficients.

The high correlation of 0·81 between tests I and II suggests that possibly the tests are so similar that they are in fact demanding much the same performance, or are at any rate overlapping to a large extent. On the surface there seems to be little in common between the two tests except that they both require good muscular co-ordination and control of movement. A possible group factor is that of 'manual skill', which may underlie both activities. Then a subject who is skilful in one activity will tend to be so in the other, and since variability is highly correlated with error, the chief reason for the high correlation of variability scores will lie in the manual factor. On the other hand, this will hardly explain the correlation with the character judgement test, in which a subject has to judge the extent to which she possesses a quality and block in a scale to that degree.

To settle the question of overlapping material, the correlation of variability scores of the arm-movement and ball-and-slot tests has been worked out with the error scores partialled out. To do this, the error results of the two tests were pooled (which the degree of correlation justified) and this pooled ranking was partialled out, giving partial coefficients of variability on the two first tests as shown in Table XI.

Table XI. *Variability—error partialled out*

		Test I (arm-movement)					
		A_1	A_2	A	B_1	B_2	B
Test II (ball-and-slot)	A_1	0·45					
	A_2		−0·22				
	A			*−0·27*			
	B_1				0·24		
	B_2					0·25	
	B						*0·37*

In this table, the correlations of the A period (0·45 and −0·22) are accounted for by the fact that in the first half the correlation of error and variability is low while in the second half it is high. Thus the variability correlations of 0·32 and 0·20, when error is partialled out, take on very divergent values because of the difference in the degree of association with the error scores.

The striking feature of this table is that, when error is partialled out between tests I and II, the coefficients come to resemble in magnitude the intercorrelations with test III in which there can be no such influence of error and skill. The results of the A periods become similarly of doubtful significance; the B results show the same low but consistent coefficients.

The conclusion to be drawn from this seems to be that some factor such as manual skill affects both the arm-movement and ball-and-slot tests and their variability scores, and that this makes these variability correlations spuriously high; that when its influence is removed, the intercorrelations of all three tests do indicate the presence of a group factor of variability manifested in the last part of the test—namely, after one half-hour work period has elapsed and the subjects are working on each test for 3 min. a day; that, however, such evidence is only tentative in view of the small number of subjects and in the absence of factor analysis.

The error intercorrelations are given in Table XII. This considerable degree of correlation suggests that throughout these two tests there is some common factor that induces success or failure in both. It is not an extremely strong tendency but it undeniably exists; and it is this that has made the variability intercorrelations artificially high.

Table XII. *Error scores*

		Test I (arm-movement)					
		A_1	A_2	A	B_1	B_2	B
Test II (ball-and-slot)	A_1	0·53					
	A_2		0·50				
	A			*0·65*			
	B_1				0·59		
	B_2					0·57	
	B						*0·60*

The intercorrelations of the speed scores are given in Table XIII. These results show that there is no general speed factor operative in these tests. This disagrees with casual observation made of subjects during the tests that some tended to work on all tests more quickly than others. It is in agreement with Vernon's (11) finding in that although there is no general speed factor common to all the tests, there is a high degree of constancy of specific speeds; i.e. subjects keep a very consistent speed in each test, as is shown by the reliability coefficients, although there is no carry-over to other tests.

Table XIII. *Speed scores*

		Test II (ball-and-slot)						Test III (character judgement)					
		A_1	A_2	A	B_1	B_2	B	A_1	A_2	A	B_1	B_2	B
Test I. (arm-movement)	A_1	0·01						0·35					
	A_2		0·05						0·07				
	A			*0·00*						*0·32*			
	B_1				0·46						−0·30		
	B_2					0·65						0·16	
	B						*0·66*						*0·01*
Test II (ball-and-slot)	A_1							−0·09					
	A_2								−0·08				
	A									*−0·09*			
	B_1										0·00		
	B_2											0·16	
	B												*0·02*

Variability has here been found to be a much more valuable source of enquiry than the other two measures taken from the work results. It has been found that subjects who are variable in one test tend to be variable in two very different tests, but only after a testing time of half an hour has elapsed. The question immediately arises, how wide is this factor? Does it extend from purely sensory judgements to judgements involving intense concentration or co-ordination of action, or are there different levels or 'fields' of variability? It would be necessary to have a larger number of subjects to answer these questions, so that each coefficient would be significant in itself, and the correspondence of variability scores in a battery of tests traced more accurately throughout the testing time.

IV. RATING SCALE RESULTS

(1) *The block-in scale* (A)

Traits in this scale were, for correlation purposes, divided into two groups:

'Persistence' traits: those traits giving the theoretical aspect of the specific situations of the hypothetical situation scale B;

'w' traits: traits of a wider classification called, provisionally, '*w*'. Two of these traits were included in the 'persistence' list and their overlapping has of course to be accounted for.

The intercorrelations of the 'persistence' traits are shown in Table XIV. In this table the correlations are nearly all positive, but 'tendency to planned action' has only a negligible correlation with other traits judged by its average correlation, and in no case does it correlate more highly than 0·31.

Table XIV

	Trait	17	24	14	11	Average
Tendency to planned action	(9)	0·27	0·31	−0·11	0·29	0·19
Permanence of interests	(17)		0·22	0·48	0·36	0·33
Thoroughness in routine work	(24)			0·36	0·50	0·35
Strict adhesion to principles	(14)				0·38	0·28
General stability of moods	(11)					0·38

The positive trend of correlation might be accounted for by the existence of a group factor of persistence or of a wider character factor such as '*w*', which would include the persistence qualities. Such factor analysis has not been attempted here, the correlations being used simply as a justification for pooling the scores on those five persistence traits and obtaining a single 'persistence' ranking. This can then be compared and correlated with the ranking obtained on the other measure of persistence, the hypothetical situation scale.

For '*w*' traits, the coefficients given in Table XV were found. It is seen immediately that trait number 27 (depth of thought) correlates negligibly with the other traits; it is hard to say why this should be, except that perhaps 'depth of thought' conveys a different meaning to the reader from 'profoundness of thought' used by Webb. In any case the coefficients are so low that this trait has been excluded, and Table XVI

Table XV

		'w' traits							
	Trait	12	13	14	24	27	28	29	Av.
Strength of will	(12)		0·14	0·33	0·15	0·22	0·27	0·35	0·24
Conventionality	(13)			0·31	0·45	−0·10	−0·07	0·34	0·18
Strict adhesion to principles	(14)				0·36	0·02	0·36	0·49	0·31
Thoroughness in routine work	(24)					−0·19	0·69	0·52	0·33
Depth of thought	(27)						−0·01	−0·08	−0·02
Conscientiousness	(28)							0·70	0·32
Trustworthiness	(29)								0·39

Table XVI

Trait	12	13	14	24	28	29	Average
12		0·14	0·33	0·15	0·27	0·35	0·25
13			0·31	0·45	−0·07	0·34	0·23
14				0·36	0·36	0·49	0·37
24					0·69	0·52	0·43
28						0·70	0·39
29							0·48

Average $r = 0·36$.

obtained. Here with one exception the coefficients are all positive, the average correlation of 0·36 being slightly higher than the average correlation for 'persistence'.

The uniformity of these coefficients allows a single '*w*' score to be made by pooling the six traits; the term '*w*' is used tentatively, because no attempt at the analysis of '*w*' has been made, the evidence from previous work on the factor being taken for granted. The intercorrelating traits bear sufficient resemblance, it is thought, to the traits described by Webb (1), Garnett (12) and others to justify the final score obtained from them being called a '*w*' score.

The correlation of 'persistence' or '*P*' (this symbol is not used to denote a general or group factor but is merely an abbreviation) with '*w*' score $= 0·77$. This, however, contains a large element of spurious correlation, since two of the traits were included in both the '*P*' and '*w*' scores. If this spurious correlation is taken as $= 2/\sqrt{30}$ (see Brown & Thomson (13)), there remains a correlation of 0·41 over and above that to be expected from the presence of the common elements. It is therefore possible that since the '*P*' traits correlate rather less among themselves than with the '*w*' scale, their intercorrelation may be due simply to the presence in all of them of some degree of '*w*' factor. This can be neither proved nor disproved in this experiment, but is merely suggested. There is certainly no direct evidence that these traits of persistence are anything more than '*w*' traits; that, in fact, the particular aspects of persistent behaviour taken here may be manifestations of 'perseverance' or 'will' or whatever else this general factor may be called.

The two scores, '*P*' and '*w*', have, however, been kept distinct for correlation with the other rating scales and finally with the test results.

(2) *The hypothetical situation scale* (B)

Contingency coefficients were obtained between three of the five questions of the scale from a 5×5 table, giving Table XVII. There is thus fair agreement between situations designed as example situations of 'persistence in routine', 'permanence of interests' and 'persistence of intentions'. At least we are justified in saying that in these eighteen cases there is a decided tendency to judge the person who would be careful at checking additions to be also persistently interested in any hobby she might take up, and again persistent in carrying through her holiday plans: and, on the other hand, the person who would neglect checking precautions would also be the person to lose enthusiasm quickly in a hobby, and to be changeable and impulsive over holiday arrangements.

Table XVII

Question		1	2	4
Persistence in routine (checking additions)	(1)		0·675	0·636
Permanence of interests (hobbies)	(2)			0·710
Persistence of intentions (keeping to holiday plans)	(4)			—

Average $C = 0{\cdot}674$ (approx. 4 times the P.E.).

Between the other two questions, in which the grading was coarser, contingency coefficients on a 2×2 table were as in Table XVIII. There seems to be little if any relation between persistence of principles and persistence of moods, and between these and the other pooled questions of the scale; at least, there is little relation between the tendency to support

Table XVIII

Question		3	5	1, 2, 4 pooled
Persistence of principles	(3)		0·330	0·219
Persistence of moods	(5)			0·00

societies with whose aims one is in sympathy and persistence of cheerful and despondent moods. On the other hand, lack of correlation may be due to the coarseness of the grading on these two questions.

There is thus some evidence—not conclusive—that persistence is not entirely specific to every activity. At least, to the minds of the judges there is a correspondence between persistence in routine work, in hobbies and in making plans. Webb's term 'perseverance in the face of obstacles' is perhaps another way of expressing this characteristic, since the example situations imply at least some resistance to impulse if not to external obstacles.

(3) *The Marston introversion-extraversion scale*

The subjects were marked on a five-point scale for each question, '1' denoting the extreme introvert score, and '5' the extreme extravert.

The total score for each rater was the sum of the marks of the twenty questions, and the final introversion score was the average of the scores awarded by the three raters.

(4) *Reliability of the scales*

The reliability of each rating scale was found by intercorrelating the scores of the three sets of judges and taking the average of the three coefficients obtained. This gave:

Block-in scale (A): '*P*'	$r=0{\cdot}82$
'*w*'	$r=0{\cdot}76$
Hypothetical situation scale	$r=0{\cdot}52$
Marston scale	$r=0{\cdot}69$

(The reliability of the Marston scale itself was found by correlating the sum of the scores on odd traits with the sum on even traits. The reliability of the half-test was then 0·83, or for the whole scale, 0·91.)

The reliability of the rating scales is unfortunately very much lower on the whole than that of the test results. In part this may be due to the usual discrepancies of ratings of personal characteristics; it may also be due to the small number of subjects and to the fact that each subject had her own judges. Had the subjects been more in number, and had they been rated by the same three judges, it is possible that the reliability would have been higher.

(5) *Intercorrelation of rating scale results*

From all the rating scales used, eight measures were finally available —four of 'persistence', one of '*w*', one of introversion and two self-ratings. The reliability of the self-ratings was found to be lower throughout than that of ratings of judges, and so these are not described in the present report.

In Table XIX the correlations given in brackets are only approximate, being the coefficients over and above those to be expected from chance alone, due to the overlapping of questions. The coefficients are now all positive, with an average correlation of 0·51. The measures are not of course entirely independent of each other, although overlapping has been ruled out as far as possible. A good degree of correspondence between 'persistence' as described abstractly and as described in specific situations was expected and hoped for; and the coefficient of 0·72 is satisfactorily high in view of the arbitrary selection of the 'hypothetical situations' and the unexplained trait descriptions of the block-in scale. The best-correlating measure is the hypothetical situation scale (average $r = 0{\cdot}64$). The lowest is the '*w*' measure, which is perhaps not surprising in view of the abstract nature of the '*w*' descriptions.

Table XIX. *Rating scale intercorrelations*

Rating scale	1	2	3	4	5	6	Average
'*P*': hypothetical situation scale (1)		0·72	0·60	0·56	0·74	0·58	0·64
'*P*': block-in scale (2)			(0·40)	0·38	0·46	0·47	0·49
'*w*': block-in scale (3)				0·35	0·50	0·28	0·43
Introversion score (4)					(0·43)	(0·49)	0·44
Question 5 of Marston scale* (5)						0·71	0·57
Question 16 of Marston scale† (6)							0·51

Average $r = 0{\cdot}51$.

* Question 5 is: "Tends towards activities requiring care; good in detail."

† Question 16 is: "Does not abandon an activity readily."

V. CORRELATION OF RATING SCALES AND TEST RESULTS

(1) *Ratings and variability scores*

The previous sections have been concerned with character ratings on the one hand—their internal consistency and intercorrelation—and with test results on the other. The main aim of the experiment was, however, to bring the two together and see how, if at all, these character traits would be revealed in accuracy, speed or variability of work.

Table XX gives the correlation of all the ratings (neglecting only the self-ratings) with the variability scores, i.e. with low variability. Considering the arm-movement and ball-and-slot tests together, it is clear

Table XX. *Variability and persistence ratings*

	Test I (arm-movement)			Test II (ball-and-slot)			Test III (character judgement)		
Rating scale	A	B	Av.	A	B	Av.	A	B	Av.
'*P*': hypothetical situation scale	−0·14	−0·04	−0·09	0·22	0·08	0·15	0·30	0·39	0·34
'*w*': block-in scale	0·07	−0·18	−0·05	−0·20	−0·30	−0·25	0·33	0·23	0·28
'*P*': block-in scale	0·25	−0·05	0·10	−0·06	−0·18	−0·12	0·24	0·19	0·21
Question 5 (Marston)	0·05	0·16	0·10	0·20	0·08	0·14	0·47	0·50	0·48
Question 16 (Marston)	0·26	0·08	0·17	0·07	−0·02	0·02	0·21	0·32	0·26
Average	0·10	−0·01	0·04	0·05	−0·07	−0·01	0·31	0·33	0·32
Introversion	−0·21	−0·18	−0·19	−0·25	−0·39	−0·32	0·17	0·16	0·16

that there is no significant relationship present; no one of the five persistence measures shows any consistent correlation with variability of work, although the negative coefficients with introversion suggest that possibly some or all of the extravert qualities tend to produce successful motor performances such as are required in these tests, and so indirectly produce low variation.

This lack of correlation was to be expected in two tests where the amount of error correlated highly with variability. It seems that the relative liability to error so masked comparative variability that the average deviation from the mean was not a true variability score, i.e. it was not an independent score, but was dependent upon the error scores. The highly inaccurate subjects, probably because of their knowledge of success and failure, had to vary their judgements more than those, more accurate, who were more satisfied with their results.

Therefore an attempt was made to unmask true variability scores to a small degree at any rate, by dividing the variability score of each subject by her error score. This was done in both the arm-movement and ball-and-slot tests. In the arm-movement test, however, the extremely inaccurate subjects then obtained the lowest relative variability scores, simply because of the size of the denominator in the ratio variability/error. The relative variability score was therefore taken as

$$\frac{\text{rank order on variability scores}}{\text{rank order on error scores}}.$$

The correlations with the rating scales were then repeated for relative variability and were as shown in Table XXI. In this table the coefficients for the continuous work of the arm-movement test are all negative, but it must be remembered that the A period of this test was

Table XXI. *Variability relative to error*

	Test I		Test II	
Rating scale	A	B	A	B
'*P*': hypothetical situation scale	−0·25	0·33	0·65	0·37
'*w*': block-in scale	−0·17	0·21	0·43	0·39
'*P*': block-in scale	−0·06	0·50	0·37	0·21
Question 5 (Marston)	−0·19	0·17	0·43	0·23
Question 16 (Marston)	0·01	0·14	0·33	0·15
Average	−0·13	0·27	0·44	0·27
Introversion	0·01	−0·10	0·32	0·47

found to be unsatisfactory—its correlation with the B period was low and it correlated poorly with variability on the other tests.

The other coefficients now show positive correlations with the rating scale results. Their actual magnitude is of little value because of the small number of subjects, but the consistently positive correlation does show that there is a tendency for lack of variability to be associated with very varying measures of 'persistence' and '*w*'—measures whose intercorrelation averages only 0·51.

It is of some interest to compare the values of the rating scales from the point of view of their correlation with the test results. The average correlation on A and B periods on all three tests has therefore been taken and is as shown in Table XXII. The best correlation is with persistence as exemplified in specific, imaginary situations. This is also the rating scale that has the best average intercorrelation with the other scales (0·64), although its internal consistency is not good (0·52). Its correlation with the test variability scores, except in the unreliable I A period, never drops below 0·29, and with the continuous work periods of the character judgement test is as high as 0·65. This suggests that the most fruitful form of rating scale to take would be a 'specific situation' scale. It would, however, be constructed in more detail than the hypothetical situation scale, and on the basis of the best intercorrelating questions.

Table XXII

	Periods		
Rating scale	A	B	Average
'*P*': hypothetical situation scale	0·47	0·36	0·41
'*w*': block-in scale	0·38	0·26	0·32
'*P*': block-in scale	0·30	0·30	0·30
Question 5 (Marston)	0·45	0·30	0·37
Question 16 (Marston)	0·27	0·20	0·23
Average	0·37	0·28	

If this were done, the internal consistency might be greatly improved, and the correlation with variability might also be higher.

The above results show that there seems to be no greater correspondence after a long time of testing than in the first half an hour's work. This is not true of the arm-movement test, in which the first half-hour seems to involve the interplay of many factors, but the ball-and-slot and character judgement tests correlate with variability equally well in the A periods of 30 min. continuous work as in the later B periods of 3 min. work a day for 10 days. It cannot be said that this would be true for other kinds of work, and it would be very hard to frame any law on the length of testing time necessary in judgement tests generally; but from the point of view of further work it would be possible to design tests such as these latter two, which need not be taken to any great length of testing time for them to give reliable variability scores.

(2) *Ratings and error scores*

The correlations of rating scales and actual error scores were obtained but found to be entirely negligible (average $r = -0{\cdot}095$). The table of correlations is therefore not given.

(3) *Ratings and speed*

The average correlation of persistence ratings with slowness was $-0{\cdot}055$, and here again the table of correlations has been omitted. Contrary to popular belief, the careful, persistent workers (i.e. those rated as such) did not prove to be any slower than those rated as more impulsive and careless.

In the intercorrelations between measures of speed on the three tests, no evidence was found for a speed factor, speed being consistent throughout a test but not overlapping with the speed scores on any other test. Here speed seems to be independent of the wide character qualities of the rating scales; so that, from a diagnostic point of view, the speed with which a subject works is no indication of what his speed will be on any other kind of work, nor is it any indication of character.

(4) *Correlation of variability with the whole block-in scale*

Variability has already been correlated with traits of the block-in scale which were pooled to form '*P*' and '*w*' estimates. The full correlation of the character judgement test (III) with all the thirty traits of the scale was also obtained, this test being chosen as that yielding variability scores uncomplicated by error. This full list includes the 'self' traits and others which have so far been left out of account. Of the thirty correlations obtained the best-correlating traits, those above 0·20, are shown in Table XXIII.

Table XXIII. *Correlation of traits with variability*

Positive correlations		Negative correlations	
Conventionality	0·62	Originality	−0·33
Strict adhesion to principles	0·29	Speed of work	−0·30
Pugnacity	0·25	Talkativeness	−0·29
Thoroughness in routine work	0·23	Imagination	−0·27
Strength of will	0·22	Quickness of understanding	−0·27
		Shyness	−0·27
Average $r = 0{\cdot}32$.		Average $r = -0{\cdot}29$.	

There are several seeming discrepancies here. In the positive correlations 'pugnacity' (in Webb's research 'readiness to become angry') is not generally associated with the '*w*' qualities, and in the negative correlations, 'shyness' and 'talkativeness' are equally correlated. But, apart from these instances, the positive correlations are on the whole with traits of a '*w*' type. The four best-correlating traits are in fact the very ones selected, quite independently and before this full correlation with the block-in scale was done, to represent '*w*'. The traits correlating negatively, on the other hand, are also easily recognizable as non-'*w*' traits—in many respects the reverse of '*w*'—'originality' and 'imagination' in place of 'conventionality'; 'quickness' and 'speed' instead of 'thoroughness'.

Thus there is evidence from several sources that variability of judgement can be measured and found to be consistent over a long-continued time of testing. There is evidence for a group factor of variability present even when the influence of accuracy has been eliminated; and last, and most important, there is evidence for a consistent, low correlation of variability with measures of 'persistence' and '*w*'. There is no information as to the closeness of relationship; the evidence shows only its existence. Such evidence rests, not upon the size of the coefficients, which would be of no value in results from a group of only eighteen subjects, but upon the general trend of the coefficients.

VI. SUMMARY OF MAIN RESULTS

1. The relations of speed, error and variability to each other were not general, but dependent upon specific conditions in each test.

2. Speed within each test was highly consistent throughout half an hour's work and the subsequent ten 3 min. daily periods, but there was no evidence for a speed factor underlying the scores on the three tests.

3. There was evidence for a factor of variability, present in all three tests, but in much greater evidence in the later 3 min. a day testing than in the earlier continuous half-hour's work.

4. The rating scales of 'persistence' and '*w*', although less reliable than the test results (as judged by the reliability coefficients) showed consistently positive correlation.

5. Low variability on the character judgement test correlated positively with all the 'persistence' and '*w*' measures at four points in the progress of testing, giving twenty coefficients in all. The small number of subjects makes the exact values of the coefficients of little value, but the general trend of positive correlation seems to show that 'steadiness' of character is associated with low variability.

6. Low variability, taken relative to the actual error scores, correlated positively with '*w*' and 'persistence' estimates in the latter half of the arm-movement test and in both types of work in the ball-and-slot test. These estimates did not, however, correlate with the crude variability scores on these tests.

7. Low variability correlated positively with introversion, although the coefficients are not high enough to be individually significant, nor sufficient in number to show the general trend of correlation.

8. The hypothetical situation scale had the best average correlation with the variability scores on all three tests. It seems therefore that the best type of rating scale is one of specific, imaginary situations rather than one listing abstract trait-names.

9. The relationship of accuracy to the 'persistence' scales was entirely negligible, though there is a slight positive correlation with extraversion.

10. There was similarly no significant relationship or trend of correlation between speed and 'persistence'.

REFERENCES

(1) Webb, E. (1912). "Character and intelligence." *Brit. J. Psychol. Mon. Supp.* I, No. 3.

(2) Cox, J. W. (1934). *Manual Skill.* Cambridge University Press.

(3) Walton, R. D. (1936). "Relations between the amplitude of oscillations in short period efficiency and steadiness of character." This *Journal*, XXVII, 181–8.

(4) Hollingworth, H. L. (1925). "Correlations of achievement within an individual." *J. Exp. Psychol.* VIII, 190–208.

(5) Asch, S. E. (1932). "An experimental study of variability in learning." *Arch. Psychol.* No. 143, 1–54.

(6) MAY, M. A. & HARTSHORNE, H. (1929). *Studies in Service and Self-control.* New York: Macmillan Press.
(7) CRUTCHER, R. (1934). "An experimental study of persistence." *J. Appl. Psychol.* XVIII, 409–17.
(8) HOWELLS, T. H. (1933). "An experimental study of persistence." *J. Abn. Soc. Psychol.* XXVIII, 14–29.
(9) GOBAT, J. D. (1937). "Error and variability of judgement in routine work; their relation to assessments of character." Thesis for Ph.D. degree. Deposited British Museum.
(10) GUILFORD, J. P. & HUNT, J. M. (1931). "Some further experimental tests of McDougall's theory of introversion-extraversion." *J. Abn. Soc. Psychol.* XXVI, 324–32.
(11) VERNON, P. E. (1933). *Studies in Expressive Movement.* London: Macmillan.
(12) GARNETT, J. C. M. (1919). "General ability, cleverness and purpose." This *Journal,* IX, 345–66.
(13) BROWN, W. & THOMSON, G. H. (1925). *The Essentials of Mental Measurement.* Cambridge University Press.

(*Manuscript received* 10 *February* 1938)

SOME RELATIONS BETWEEN REPRODUCING AND RECOGNIZING PROSE MATERIAL

By O. L. ZANGWILL
Cambridge Psychological Laboratory

I. *Introduction* (pp. 370–371).
II. *Experimental method* (pp. 371–376).
  (1) *Material and instructions* (pp. 371–372).
  (2) *Conditions of reproduction* (p. 372).
  (3) *Material for recognition* (pp. 372–375).
  (4) *Instructions for recognition* (pp. 375–376).
  (5) *Subjects* (p. 376).
III. *Principal results* (pp. 376–378).
IV. *Some qualitative observations* (pp. 378–380).
  (1) *Construction and style in relation to recognition* (p. 378).
  (2) *Familiarity and unfamiliarity* (p. 378).
  (3) *The dominance of details* (pp. 379–380).
V. *Positive and negative relations in recognition and recall* (pp. 380–381).
VI. *Summary and conclusions* (pp. 381–382).
References (p. 382).

I. INTRODUCTION

IN an earlier paper (3), some evidence was produced to the effect that the various varieties of constructive and dynamic processes which have been shown to operate in remembering are not without influence on recognition. This evidence was neither certain nor conclusive; yet the relations which we were able to establish between some aspects of recognition and recall seemed of sufficient interest to justify further study.[1]

The present investigation departs from its predecessor in two important respects. In the first place, connected prose passages were used instead of simple figures; in the second place, the conditions were arranged with a view to minimizing the effects upon recognition of interpolated recall. The technique of repeated reproduction was entirely discarded, and the conditions were, in consequence, better adapted to the study of anomalies in the process of recognition itself.

---

[1] It is perhaps worth remarking that Hanawalt's extensive results (2) have largely confirmed the modest observations recorded in our preliminary study.

The general plan of the experiment was identical with that of the earlier investigation. A number of untrained subjects were instructed to read through an unfamiliar prose passage and to make no attempt to memorize the material. Three months later, a reproduction of this passage was requested from each of the subjects. Unknown to the latter, the material of these reproductions was utilized by the experimenter in the preparation of special passages for recognition. Some time later, any subject who had furnished a reasonably detailed recall was provided with three superficially similar prose passages. These comprised (*a*) the original narrative; (*b*) a passage of approximately equal length and not altogether unlike the former in style and content; and (*c*) a narrative constructed by the experimenter in such a way as to include the most salient inaccuracies of the individual's own reproduction. The various transformations and inventions which had originated in recall were welded together into the form of a coherent narrative bearing a strong resemblance to the original in style and construction. Each subject was instructed to read through the three passages, and to indicate in a prescribed manner the one which he believed to be the original. He was also asked to specify the certainty of his judgement and the factors upon which he believed it to be based.

II. EXPERIMENTAL METHOD

(1) *Material and instructions*

The passage selected was a portion of a Hindu legend reproduced by Oesterreich in his book on *Possession*. It was chosen as a suitable vehicle for the general tendencies of 'conventionalization' whose operation in recall has been so strikingly illustrated in the studies of Bartlett. The narrative reads as follows:

> There is a town of the name of Vatsamân, where lived a Brahman called Kecava who was wise, but poor. His wife, who was called Karagarâ (i.e. poison-giver), behaved so ill towards everyone that even a demon, who lived on a tree in the house, fled into the desert for fear of her. Meanwhile the Brahman was no longer able to bear the wickedness of his wife and went away also. On the way to the desert the demon saw him and said to him: "I wish to-day to offer thee hospitality." When the Brahman heard this he was afraid. "Fear not," said the demon, "for I used to live upon a tree in thy house, but I fled to this spot for fear of Karagarâ, and since thou hast long had to do with me as landlord I will do good to thee. Go thy way to the village of Mrigavatî (that is to say, rich in Gazelles); I will take possession of the daughter of the city, Mrigalotschanâ (signifying gazelle-eyed), and will not let myself be driven out by any exorcist; thou alone, when thou comest, shall drive me out with a look." Having spoken thus the demon entered into the royal virgin. Meanwhile the Brahman went

to the royal city Mrigavatî and having heard the herald he went to the royal palace, but in vain he did all that magicians are accustomed to do and uttered his conjurations, the demon did not leave the maiden. When the Brahman saw that in no other way would the demon come out, he cried: "In the name of Karagarâ, come forth!" The demon replied: "See, I am coming forthwith!" and he immediately came out. Then the King gave to the Brahman the half of his kingdom and his daughter in marriage.

Each subject received a copy of the above story together with the following instructions:

> Please read through the enclosed story twice, at your normal rate of reading. As soon as you have thus read it, please destroy it *without fail.* You will not be asked to reproduce the story at any stage of the experiment, *so please make no attempt to learn the whole or parts of it by heart.*
>
> Further instructions will be forwarded to you in due course.

The entire experiment was carried out without direct contact between the experimenter and the subjects, all communication being reciprocally postal. Thus although the requisite control of the conditions depended entirely upon the goodwill and co-operation of the subjects, the somewhat artificial atmosphere of the usual laboratory experiment was largely avoided.

(2) *Conditions of reproduction*

After a lapse of three months, the following communication was sent to the subjects:

> Some time ago, you received a short story which you were asked to read and then to destroy. We should be grateful if you could write down on the enclosed paper as much of this story as you can remember and forward it to us in the enclosed envelope. If you are unable to remember the story coherently, please write down any detail or indeed anything connected with the story that you may be able to recall. Should you be unable to remember anything whatsoever, please return the enclosed paper stating this fact.
>
> We should like to apologize for our apparent infraction of the original instructions, but it was essential for our purposes that intentions to remember should be rigidly excluded.
>
> The instructions for the final stage of this experiment will be forwarded to you very shortly.

(3) *Material for recognition*

The mode of construction of the material provided for recognition has already been mentioned. It will be remembered that each subject received three passages, labelled A, B and C respectively. Version C was the original story; version B a specially prepared narrative based on the reproduction; and version A an unfamiliar narrative introduced as a kind of control. For purposes of comparison, the narrative A is reproduced below:

> There is a town of the name of Vatsamân where lived a Brahman called Kecava who was wise but poor. The good Brahman and his family were greatly tormented by a demon, who lived upon a tree in the house. Earnestly the Brahman pleaded with the demon, and finally the latter agreed to obey his bidding and departed from this place. When the demon had come forth from the tree, he went into the town of Karnavatî (that is to say, the town with ears) and took possession of the queen, who was half-sister on the father's side of the King Madana and who was called Sulotschanâ, (signifying lovely-eyed). Greatly tormented by the demon, the queen became like a skeleton. Then the king, whose name was Satrughna (slayer of enemies), sent to the King Madana and begged him to send the wise Brahman Kecava, who had rid the town Vatsamân of a demon who lived in a tree. At the request of Madana, Kecava came to Karnavatî to the possessed queen. But when the demon perceived him he insulted and threatened him: "It is enough that I have done thy bidding once; now take care and look to thyself!" When the Brahman heard this, he recognised that it was the same demon. When the Brahman saw that in no other way would the demon come forth, he uttered his magical conjurations. The demon was seized with fright and immediately left the queen. The Brahman, covered with honours by the King, made his way back to Mrigavatî.

Seven of the nineteen subjects gave recalls sufficiently detailed and coherent to furnish the basis of special recognition versions. One of the seven indeed provided sufficient material for the construction of two such versions (in this case labelled B and D respectively).

As illustrations of the method used in construction of the recognition material, two complete examples may be given:

Subject K: Recall version. I remember the scene of the story as being set in the Kashmere-Punjab district in a village set in well-wooded foothills. Its name began with M. There was a middle-aged man of rustic profession (e.g. woodcutting?) therein, with a horrid mother-in-law (or step-mother). Both their names began with M. From the forests near by a were-tiger (or maybe just an ordinary tiger) made depredations on the villagers and their gardens. The wood-cutter somehow killed it. At this point the step-mother seems to go out of the picture. Perhaps she was identified with the tiger. The next scene is the charcoal burner (*sic*) leaving for some reason (the old mother's malign influence?) and going south into the environs of an important city of the Plains, where he met a princess, whom however I cannot remember him actually marrying, though that seemed to be the natural outcome if the story had not ended before it could happen; her name began with an M.

Recognition (B) version. There is a village of the name of Vatsamân which is surrounded by hills and woodland. Here lived Kecava, who though but a poor wood-cutter, was a Brahman and very wise. Kecava and his wife suffered greatly under Karagarâ (i.e. poison-giver), mother-in-law to Kecava, who behaved so ill towards everyone that even the leaves of a tree in the courtyard withered and no flowers would bloom. At that time, the villagers were suffering greatly from the depredations of a tiger who stole forth at nightfall from the forests, and whose ravages struck terror into the hearts of the humble villagers. One evening, the Chief of the village Elders approached Kecava and said to him: "Never before has mortal beast so ravaged our peaceable villagers. Our Elders have held council, and believe this to be the work of the evil woman Karagarâ, the Giver of Poison. It is she who takes possession of the tiger and leads it to ravage our village." Kecava went forth, and after many weeks of search, he encountered the tiger, that had ravaged the village of Vatsamân. A great struggle took place, and the tiger was slain by the hand of the Brahman. The wanderings of Kecava had carried him south into the plains and soon he reached the Royal City of Mrigavatî. The people of this city were in mourning because the Royal Virgin, Mrigalotschanâ (which signifies gazelle-eyed) had been possessed by a demon. Having heard the herald, Kecava proceeded to the Royal Palace, but in vain did he do all that magicians are accustomed to do; the demon did not leave the maiden. When the Brahman saw that in no other way would the demon come forth, he cried: "In the name of Karagarâ, come forth!" Immediately the demon came out and the maiden became once more her true self. Then the king gave to the Brahman the half of his kingdom and his daughter in marriage.

Subject O: Recall version. A Brahmin was having supper one night with his wife when he heard a loud wailing and screaming in the garden. Rushing out, they discovered a demon caught in the branches of a pomegranate tree from which he was unable to extricate himself. The woman was terrified and rushed into the house again, but the Brahmin stood his ground and held conversation with the demon. In the course of this the spirit promised that if he were released from the tree so that he could return to his own works, he would be at any time at the order of the Brahmin. At this the Brahmin fetched a ladder, cut away with an axe the encircling branches, and the demon flew away.

Shortly after this the Brahmin's wife died and he departed from the place in which he lived. He walked for days through a desert until at last he came to a city which was called "The Pools of the Gazelles", and which was ruled by a King the name of whose daughter, the Princess, translated means "Gazelle-eyed". The people of that city were in mourning because the princess was possessed by a demon, and although two other famous sorcerers were there, their efforts to dislodge the demon were in vain. The Brahmin approached the princess and said in a loud voice "Come forth!" "I come", replied a voice which was the voice of the spirit which had been caught in the pomegranate tree. The princess was immediately perfectly normal.

Recognition (B) version. There is a town of the name of Vatsamân, where lived a Brahman called Kecava who was wise but poor. His wife, who was called Karagarâ (i.e. the poison-giver), behaved so ill towards everyone that even a demon who lived on a pomegranate tree in the garden, flew away wailing and screaming into the desert for fear of her and returned to his own works. Meanwhile the Brahman was no longer able to bear the wickedness of his wife and went away also. The demon met him and said to him: "I wish to offer thee hospitality." When the Brahman heard this he was afraid. "Fear not," said the demon, "for I used to live upon a pomegranate tree in thy house, but I fled to this spot for fear of Karagarâ. And since thou hast long had to do with me as my landlord, I will do good to thee. "Go thy way to the city Mrigavatî (that is to say, 'The Pools of the Gazelles'); I will take possession of the daughter of the city, Mrigalotschanâ (signifying Gazelle-eyed), and will not let myself be driven out by any exorcist. Thou alone shalt drive me out with a look when thou comest." Having spoken thus, the demon entered the Royal Virgin. Meanwhile the Brahman walked for many days through a desert until at last he came to a city which was indeed Mrigavatî. The people of this city were in mourning, because the princess was possessed of a demon, and although two famous sorcerers were there, their efforts to dislodge the demon were in vain. Then the Brahman approached the princess and cried in a loud voice: "In the name of Karagarâ, come forth!". "I come!" replied a voice which was the voice of the spirit who had lived in the pomegranate tree. The princess was immediately perfectly normal. Then the King gave to the Brahman the half of his kingdom and his daughter in marriage.

Recognition (D) version. There is a town of the name of Vatsamân, where lived a Brahman called Kecava who was wise but poor. As Kecava and his wife, called Karagarâ, were sitting one night at their supper, they heard a terrible wailing and screaming. Rushing out from the house, they discovered a demon caught among the spreading branches of a large tree which grew in the courtyard. Karagarâ was overcome with terror, and fled back trembling into the house, but the wise Brahman approached the demon and held conversation with him. "If thou releasest me from the imprisoning branches of this tree," pleaded the demon, "then will I promise to serve thee faithfully and to be ready at thy service whenever thou shouldst need me." The Brahman then cut away the encircling branches and the demon flew away. Shortly after this, the wife of Kecava perished, and the good Brahman departed from the place Vatsamân in which he had lived so long. On his way to the desert, he passed through the royal city of Mrigavatî (that is to say, rich in gazelles). Here he found the Royal Virgin, Mrigalotschanâ (signifying gazelle-eyed), to be possessed by a demon. Having heard the herald, he went to the royal palace, but in vain he did all that magicians are accustomed to do and uttered his conjurations; the demon did not leave the maiden. When the Brahman saw that in no other way would the demon come out, he cried in a great voice: "Come forth!" "See, I am coming forthwith", replied a voice which was the voice of the demon that had been imprisoned in the Brahman's tree. Immediately he came forth, and the Royal Virgin became once more herself.

The above examples furnish excellent illustrations of constructive recall. In the case of Subject O, it will be seen that the more significant deviations from the original narrative (18 in all) are included in version D, whereas the relatively minor inaccuracies of detail have been introduced into version B.

(4) *Instructions for recognition*

Between two and three weeks after the recall versions had been submitted, each of the aforementioned seven subjects received the following instructions:

Enclosed are a number of stories, *one* of which you were asked to read as the initial stage of the experiment. You are requested to attempt to recognize (i.e. to identify) the original story. Each story is marked A, B or C. Please place a cross on the enclosed card against the letter corresponding to that of the story which you believe to be the original. Please indicate your degree of certainty by writing 'certain', 'probably', 'possibly' (as the case may be) alongside the cross. Any introspective reports about the processes that culminated in your verdict will be welcome.

(5) *Subjects*

The subjects comprised nineteen highly educated adults of both sexes and of diverse age and occupation. Of the seven subjects who received recognition versions, five were women and two men. None of the subjects had at any time followed courses in experimental psychology and all but one were unfamiliar with Bartlett's work on remembering.

III. PRINCIPAL RESULTS

The general outcome of our inquiry is best presented in the following manner. In the present section, certain outstanding results will be given in tabular form. In the following section, a number of qualitative observations relating to the general conditions of recognition will be discussed. Finally, in the penultimate section, we shall attempt a comparison of our results with those of the earlier investigation.

The principal results are as follows:

(1) The original story (recognition version C) was identified with unqualified conviction by only one of the seven subjects to whom recognition material was submitted.

(2) Two subjects identified the original story with a fairly high degree of certainty. Both of these nevertheless admitted that version B (that based upon the respective reproduction) might just possibly be the original.

(3) Three subjects affirmed that none of the three versions was the original. Two of these three actually expressed the belief that the original story had included passages from both the B and the C versions.

> I imagine the stories B and C both contain parts of the original one. *Subject H.*
>
> Finally I said the story I saw was C with the first third of B worked into it but here left out. In other words, none of these three is the exact story. *Subject K.*

The third subject left her card blank, but her introspective report furnishes an admirable example of the reconstructive character of recognition under these conditions. It may well be quoted intact:

> None of these stories is the original story which I saw, although D is the nearest version of it. Yet the lame ending of D proves that that version has been tampered with. A has a different characterization from the original story, and I most certainly never saw it before. B has the pomegranate tree which I remember and the correct translation of the city's name, but the original story was not as long as B. As a matter

of fact, when I look at these again, I begin to believe that the original version was C, and that I chose D because it was likest to my own version. *Subject O.*

Another subject, though ultimately deciding in favour of C, wrote the following note in her report:

> If I didn't believe in the integrity of the compiler of this test, I should think that none of these three is the original. As it is, the nearest is C.

(4) All subjects were convinced (though on very different grounds) that the control story (A) did not constitute the original.

(5) It is interesting to consider the nature of the verdict in relation to the extent of transformation in recall. It must however be clearly understood that the 'points of deviation' listed in Table I are not to be treated as psychologically comparable. In some cases, a recorded 'deviation' consists only in some transformed, or imported, detail; but in other cases, a single 'deviation' may automatically involve widespread change in plot and sequence. Deviations of the latter type had of course to be taken into account in constructing the recognition material and they frequently gave rise to very striking distortions of the original narrative.

Table I

Subject	Points of deviation in recall introduced into recognition material	Nature of the verdict
D	2	C, probably; B, just possibly
M	3	C, certain
R	3	C, almost certain; B, just possibly
H	6	Verdict initially in favour of C, but original finally believed to embody parts both of B and C
V	9	Verdict ultimately in favour of C, but tendency to treat original as embodying parts both of B and C
K	9	No final positive verdict on any story. Original treated as a blend of B and C
O	15 (B) 18 (D)	No final positive verdict on any story. Considerable uncertainty and confusion in regard to versions B, C and D

Table I indicates clearly that the nature of the verdict varies significantly with the number of deviations manifested in recall and introduced by the experimenter into the recognition material. We are able in fact to conclude that where reproduction is highly constructive, recognition of the original passage under our experimental conditions is likely to become both difficult and uncertain. Moreover, there is evidence of a strong tendency in some of our subjects to treat the original as a blend. They affirm that certain passages in the B versions were present in the original beyond all reasonable doubt. This naturally gives rise to the suspicion that none of the three versions is the original.

IV. SOME QUALITATIVE OBSERVATIONS

(1) *Construction and style in relation to recognition*

Bartlett ((1), p. 81) has pointed out that the presentation of complex verbal material tends to evoke immediate and characteristic responses to general construction, syntax, rhythm and style. These factors are, he finds, extremely readily responded to, but extremely rarely reproduced with any fidelity. Bartlett's contention was amply borne out in the present study. In particular, it was observed that extremely few of the subjects made any serious attempt to recapture the style of the original narrative in their reproductions. But it is important to note that an impression of style may nevertheless persist, and be of significance in directing recognition. This is well shown in the following extract from the report of subject D:

> Much troubled by the dog in B, but C seems to have the shape and style of the original rather more. (The dog mentioned in the report was an error in recall inserted by the experimenter into version B.)

In general, it seems that responses to style and construction, while of little importance for reproduction, are not without significance in relation to recognition. This conclusion is certainly in accord with the experience of the skilled literary critic.

(2) *Familiarity and unfamiliarity*

Feelings of familiarity or unfamiliarity were seldom mentioned by our subjects in their reports on recognition. In a few cases, isolated details, or phrases in the dialogue, were said to call out immediate reactions of familiarity. But in no single case was the narrative as a whole said to give rise to an experience of this character. In general, our results lend little support to the view that recognition invariably depends upon the appearance in consciousness of an 'ultimate and irreducible' feeling of familiarity.

(3) *The dominance of details*

Bartlett ((1), p. 31) has pointed out that all perceiving involves a kind of unwitting analysis. Subjective and objective factors together combine to produce a diversification of emphasis in the perceptual field. There are always some features of the perceptual situation which take a lead over the others. There are always some details likely to stand out and influence what is perceived more than the rest. This distribution of emphasis within a perceptual field is the phenomenon of *dominance.*

The significance of dominance is by no means restricted to perception in the narrow sense. The phenomenon plays an important part in the shaping of responses remote in time from the original perceptual activity. "These dominant details", writes Bartlett, "are a kind of nucleus about which the rest cluster. They set the stage for remembering."

Our present investigation provides some data bearing upon the phenomenon of dominance and its implications for recognition. In the first place, it was found that outstanding detail appears to provide the principal clue in identifying prose material under our experimental conditions. Indeed the outcome of a judgement upon any given passage seems to be almost entirely a function of these responses to detail. The recognition of a striking detail appears to carry with it an overwhelming conviction that the entire passage must be familiar.

Some examples of the dependence of the verdict upon responses to detail may now be given:

(1) The presence in version C of the term "poison-giver" and of the phrase "the demon entered into the royal virgin" led Subject H to an immediate positive verdict upon this passage. Her verdict was subsequently qualified only through the inclusion in version B of a detail, which though absent in the original, nevertheless appeared to acquire a certain relative dominance. This is discussed below.

(2) The final selection of version C by Subject V was principally due to a phrase in the dialogue which occurs in this version alone.

A second finding of some interest is that detail introduced by a subject in remembering may apparently function as dominant and significantly influence the process of recognition. Indeed the acceptance of an invented detail as genuine was the most frequent cause of error in discrimination. This is illustrated by the following extracts from individual reports:

(1) Version B has the pomegranate tree which I remember and the correct translation of the city's name. *Subject O.*

At no point in the original was there any mention of a "pomegranate tree" and the alleged "correct translation" of the name of the city itself originated as an inaccuracy in recall.

(2) A point that also remained in my mind strongly in the original was the description of "the leaves of the tree in the courtyard withering and no flowers blooming". Well, that sentence comes in B, so I imagine the stories B and C both contain parts of the original one. *Subject H.*

The passage in question does not occur in the original narrative.

(3) I remember definitely the words: "In the name of Karagarâ, come forth!" On the other hand, I *did* think he was a woodcutter. If I didn't believe in the integrity of the compiler of this test, I should think that none of these three is the original. *Subject V.*

The substitution of 'woodcutter' for 'Brahman' characterized the reproduction of this subject.

(4) On the negative assumption that I should have remembered one or two phrases in *B*, e.g. "The voice was the voice of the traveller", I decided it was C. *Subject R.*

The last extract furnishes a good example of the negative implications of this principle of dominance. The subject in question is obviously aware of the existence of a selective process in perception and recall.

The above illustrations enable us to conclude that detail introduced into the narrative in recall may operate as dominant in much the same manner as outstanding material genuinely present in the original. This conclusion might have an important bearing upon questions of testimony and paramnesia in daily life.

V. POSITIVE AND NEGATIVE RELATIONS IN RECOGNITION AND RECALL

In our earlier investigation, positive relations[1] were found between persistent deviations in a given series of reproductions and the corresponding characteristics of a figure recognized as the original. Positive relationships are not universal, but were observed in 62 % of the total classified errors. It was further remarked that both positive and negative relationships might sometimes be demonstrated within the same objective figure.

If we take into account the vast differences in the nature of the material, in the conditions of presentation and reproduction, and in the duration of the intervals between the various stages of the experiment, it must be admitted that the results of the present research are in remarkable agreement with those of the earlier investigation. In both cases, deviations originating in recall are treated as genuine characteristics of the original and are recognised as such with a high degree of certainty. In other words, positive relationships have now been demonstrated not only in the case of relatively meaningless diagrams, but also in the case of complex and meaningful prose material. It has also been shown that some of the deviations are not treated as characteristics of the original narrative. These are examples of negative relationships very similar in kind to those reported in the earlier investigation. Lastly, it has been shown that the existence both of positive and negative relations within a given passage gives rise to considerable difficulty in expressing a judgement of recognition upon the passage as a whole. Precisely similar difficulties were encountered by the subjects in the earlier study.

---

[1] The expressions 'positive' and 'negative' relation have been defined elsewhere by the writer ((3), p. 266). They have been adopted, with some modification, by Hanawalt ((2), p. 68).

VI. SUMMARY AND CONCLUSIONS

The investigation reported in this paper represents an attempt to study some interrelationships between processes of reproducing and recognizing prose material. The principal conclusions are as follows:

(1) Recognition of a prose passage may be rendered exceedingly difficult by submitting it together with specially prepared material based upon a reproduction of the original previously furnished by the subject. In some cases, recognition seems to be virtually impossible; in others, there is a marked tendency to conclude that the original passage must have included a number of the fictitious points which originated in recall. It is suggested that these observations may have some bearing upon paramnesia in daily life.

(2) Factors of style and construction, though rarely reproduced with any accuracy, are nevertheless influential as cues in selective recognition.

(3) Recognition is frequently based almost entirely upon the response to some outstanding detail. Errors in recognition for the most part derive from similar responses to importations and constructions. The significance for recognition of 'dominant' detail is thus fairly considerable.

(4) Distinct feelings of familiarity are seldom reported. Where they occur, they are said to be aroused by relatively isolated details and not by the original passage as an undivided whole.

(5) It is pointed out that the results of this investigation parallel very closely those of an earlier study in which simple designs were used as the

experimental material. Both positive and negative relationships are found to exist between the recognition and the recall of narrative prose.

I should like to express my thanks to Mr R. N. Sinha for some help in the selection and preparation of the material.

REFERENCES

(1) BARTLETT, F. C. (1932). *Remembering.* Cambridge University Press.

(2) HANAWALT, N. G. (1937). "Memory trace for figures in recall and recognition." *Arch. Psychol., N.Y.*, no. 216.

(3) ZANGWILL, O. L. (1937). "An investigation of the relationship between the processes of reproducing and recognising simple figures, with special reference to Koffka's trace theory." *Brit. J. Psychol.* XXVII, 250–76.

(*Manuscript received* 6 *June* 1936)

THE EFFECTS OF ERRORS OF MEASUREMENT ON CORRELATION COEFFICIENTS

By ROBERT H. THOULESS

I. *The correction for attenuation* (pp. 383–387).
II. *Objections to correction for attenuation and difficulties in its use* (pp. 387–391).
III. *Conditions in which coefficients of correlation should and should not be corrected for attenuation* (pp. 391–396).
IV. *Partial correlation* (pp. 396–400).
V. *Summary* (pp. 400–401).
Appendix (pp. 401–402).
References (p. 403).

I. THE CORRECTION FOR ATTENUATION

It was first pointed out by Spearman (1) that, since no psychological test is perfectly self-consistent, there is a constant tendency for test intercorrelations to be lower than they would be if they had been obtained from self-consistent test results. This lowering of test intercorrelations by errors of measurement in the tests was called by Spearman 'attenuation', and he showed how an estimate of the 'true' or 'unattenuated' coefficient of correlation might be obtained by the use of measures of self-consistency or 'reliability' obtained by correlating two independent measurements made on the same group of persons by each of the two tests in question.

The independent measurements necessary for this purpose may be obtained in any one of three ways, the choice between them being determined by considerations of applicability of the methods or their convenience for the particular tests used. The methods are:

(1) Repetition of the same tests either on the same day or on some different day.

(2) Division of the results of each test into two equivalent halves, e.g. by adding together separately the scores for the odd and for the even items in each test.

(3) Use of two pairs of tests, of which the two members of each pair are of the same form, the same kind of material and the same difficulty, but contain different material, e.g. the L and M forms of the new Terman Revised Intelligence Scale.

The method of correction is the same in all cases. Formulae of two types were suggested by Spearman. In the present discussion, I shall confine myself to the first of these (which has been most generally used) and to the simplest case in which there are only two repetitions of each test (which also has been the usual condition in which the correction has been used).

The formula in question, as given in Spearman's first article, is

$$r_{AB}=\frac{\frac{1}{4}\,(r_{A_1B_1}+r_{A_2B_2}+r_{A_1B_2}+r_{A_2B_1})}{\sqrt{(r_{A_1A_2}\cdot r_{B_1B_2})}}.$$

Although the numerator of this equation is here given as the arithmetic mean of the four intercorrelations of the A's with the B's, the proof given by Spearman himself in a later article (2) and that supplied by Udny Yule (published as an appendix in a still later article by Spearman (3)) both lead to the geometric mean of these quantities. Spearman says in his first article that the geometric mean should be used if one test is much more accurate than the other. Although it will not make much difference which form is used if the r_{AB}'s are all about equal, the geometric mean seems to be the more correct and is always given by Spearman in his more recent writings. The formula then becomes

$$r_{AB}=\frac{\sqrt[4]{(r_{A_1B_1}\cdot r_{A_1B_2}\cdot r_{A_2B_1}\cdot r_{A_2B_2})}}{\sqrt{(r_{A_1A_2}\cdot r_{B_1B_2})}}.$$

In using this formula, it is necessary, of course, to know its standard error. This, as given by Kelley (4), is

$$\frac{r_{AB}}{2\sqrt{N}}\left(4r_{AB}{}^2+\frac{1}{r_{AB}{}^2}+\frac{1+r_{A_1A_2}+r_{B_1B_2}}{\sqrt{(r_{A_1B_1}\cdot r_{A_2B_2}\cdot r_{A_1B_2}\cdot r_{A_2B_1})}}\right)+\frac{1}{r_{A_1A_2}{}^2}+\frac{1}{r_{B_1B_2}{}^2}$$
$$-\frac{4}{r_{A_1A_2}}-\frac{4}{r_{B_1B_2}}-2\Big)^{\frac{1}{2}}.$$

The simplest proof of the formula for correction for attenuation is the one first given by Udny Yule (3), which depends on the formula for the correction of sums. If a test X is administered on two occasions and the two sets of results are not perfectly correlated, we can consider the deviations of the first test as made up of two parts, one perfectly correlated with a corresponding part of the second test deviations and the other independent of the deviations in the second test. That is, $X_1=x+e_1$ and $X_2=x+e_2$, where x is the part common to the deviations of the first and second applications of the test and e_1 and e_2 are the parts of the deviations which show independence of each other and of x. Similarly, we may use the symbols y, f_1 and f_2 for the corresponding parts of the deviations of Y_1 and Y_2 respectively.

Now $\Sigma X_1 . X_2 = \Sigma\,(x+e_1)\,(x+e_2) = \Sigma x^2 + \Sigma x . e_1 + \Sigma x . e_2 + \Sigma e_1 . e_2$. But $\Sigma X_1 . X_2 = r_{X_1X_2} . \sigma_{X_1} . \sigma_{X_2}$, and similarly for the terms on the right-hand side of the equation. So the above equation may be written

$$r_{X_1X_2} . \sigma_{X_1} . \sigma_{X_2} = r_{xx} . \sigma_x^2 + r_{xe_1} . \sigma_x . \sigma_{e_1} + r_{xe_2} . \sigma_x . \sigma_{e_2} + r_{e_1e_2} . \sigma_{e_1} . \sigma_{e_2} .$$

But since $r_{xx} = 1$, and by definition r_{xe_1}, r_{xe_2} and $r_{e_1e_2}$ are zero, it follows that

(*a*) $r_{X_1X_2} . \sigma_{X_1} . \sigma_{X_2} = \sigma_x^2$ and $r_{Y_1Y_2} . \sigma_{Y_1} . \sigma_{Y_2} = \sigma_y^2$.

By a similar argument:

(*b*) $r_{X_1Y_1} . \sigma_{X_1} . \sigma_{Y_1} = r_{X_2Y_2} . \sigma_{X_2} . \sigma_{Y_2} = r_{X_1Y_2} . \sigma_{X_1} . \sigma_{Y_2} = r_{X_2Y_1} . \sigma_{X_2} . \sigma_{Y_1}$
$= r_{xy} . \sigma_x . \sigma_y$.

From which we can easily derive the relationship: r_{xy} (the correlation between X and Y if there are no errors of measurement) is equal to

$$\sqrt{\frac{r_{X_1Y_1} . r_{X_2Y_2}}{r_{X_1X_2} . r_{Y_1Y_2}}} \quad \text{and also to} \quad \sqrt{\frac{r_{X_1Y_2} . r_{X_2Y_1}}{r_{X_1X_2} . r_{Y_1Y_2}}}.$$

Spearman's formula as given above is plainly got by taking the geometric mean of these two equations. This combination is sometimes legitimate and sometimes it is not; a better estimate of the unattenuated coefficient will, in certain cases, be obtained by taking it as equal to

$$\sqrt{\frac{r_{X_1Y_2} . r_{X_2Y_1}}{r_{X_1X_2} . r_{Y_1Y_2}}}.$$

It is important to notice, in the proof of this formula, that only the two equations (*a*) are mathematical identities. We have made these so by defining e_1 and e_2 as the parts of the deviations of the two testings of X which are independent of each other and of x. The four equations (*b*) are not mathematical identities, since we cannot define e_1 as the part of the deviations of X_1 which are uncorrelated with X_2 and also as the part uncorrelated with Y_1 and Y_2. This limitation is lost sight of if we begin by defining e_1, e_2, etc. as the 'errors of measurement' of the tests to which they belong.

Two consequences follow from this fact. First, the obvious consequence that the quantities, r_{ye_2}, etc., which are assumed to be zero in deriving the (*b*) set of equations will not be exactly zero but will at best approximate to zero within limits determined by the amount of their sampling error. This effect is, of course, allowed for in the calculation of the standard error of the final corrected coefficient. Its most striking consequence is that it may give the corrected coefficient a value greater than unity.

A second consequence is that some of the relationships indicated in the (*b*) set of equations may be altogether false if there is some special condition in the way of giving the tests which causes significant correlation between e_1 or e_2 and f_1 and f_2. There is an obvious way in which this might occur. Let us suppose that X_1 and Y_1 are administered on one day and that X_2 and Y_2 are administered on some other day, and that the mental function measured by X and Y is not constant in amount but varies for each subject from day to day. Then $r_{e_1f_1}$ and $r_{e_2f_2}$ will not be zero but will have a positive value which will depend on the extent to which any individual's amount of the function measured depends on the time of the measurement. $r_{X_1Y_1}$ and $r_{X_2Y_2}$ will, therefore, be greater than $r_{X_1Y_2}$ and $r_{X_2Y_1}$, and a true correction for attenuation cannot be obtained from $r_{X_1Y_1}$ and $r_{X_2Y_2}$, since the factor of contemporaneity which enhances these coefficients is not present in $r_{X_1X_2}$ and $r_{Y_1Y_2}$ in the denominator of the corrected coefficient (5). The formula to be used in correction for attenuation must, therefore, be $\sqrt{\frac{r_{X_1Y_2} \cdot r_{X_2Y_1}}{r_{X_1X_2} \cdot r_{Y_1Y_2}}}$.

Since this use of the two 'different time' test intercorrelations sufficiently determines the value of the corrected coefficient (whether or not $r_{X_1Y_1}$ and $r_{X_2Y_2}$ are greater than $r_{X_1Y_2}$ and $r_{X_2Y_1}$), it is suggested that this formula should always be used when X_1 and Y_1 are measured together at a time different from that of X_2 and Y_2. Otherwise, the use of the formula involving all four intercorrelations seems unobjectionable. It should be noticed that the above formula gives the correlation which would have been found between X and Y if neither errors of measurement nor fluctuation of the function measured had been present, i.e. the effect of function fluctuation in lowering the coefficient of correlation between X and Y has been eliminated as well as the effect of errors of measurement. The use of the ordinary formula would, on the other hand, have given a spuriously high value of the corrected coefficient if there had been day-to-day fluctuation of the mental function measured, and it might have given a corrected coefficient significantly greater than unity. It seems not unlikely that when correlation coefficients corrected for attenuation have been found to be significantly greater than unity, the error has arisen in this way. Fluctuation of functions measured by tests is apparently not uncommon when these functions are character traits (5). It appears to be absent in intelligence, and this may apply also to any measurement of an ability.

It seems that this same cause might result in a spurious lowering of the corrected coefficient under another condition of experiment. Let us suppose that test X was applied on one day and test Y on another, and the method of equivalent halves was used for correcting attenuation. Then $r_{X_1X_2}$ and $r_{Y_1Y_2}$ would be increased (if function fluctuation were present) by having been measured at the same time, but this increase would not be found in any of the four correlations between X and Y. Thus the effect of function fluctuation will be to increase the denominator of the correction formula while leaving unaffected the numerator. There seems to be no easy way of evading this difficulty, so that this arrangement of the tests would appear to be one to be avoided unless it is quite certain that we are dealing with a function which does not show day-to-day fluctuation.

II. OBJECTIONS TO CORRECTION FOR ATTENUATION AND DIFFICULTIES IN ITS USE

The attitude of psychological testers towards correction for attenuation has been very variable. A few seem to have used it generally as a routine treatment of correlation coefficients; a few have rejected the method either because they regarded it as inappropriate to psychological data or because they were rendered uneasy by an abnormality of corrected coefficients which was criticized by Karl Pearson; probably the largest number have admitted the necessity for correction in theory but ignored it in practice. The suggestion I am putting forward in the present paper is that correction for attenuation is essential in a certain number of problems for which correlation coefficients are used, that it is unnecessary in a much larger number of problems, and definitely incorrect for certain purposes.

Before trying to distinguish the purposes for which correction for attenuation is necessary, is unnecessary, and is wholly inappropriate, it may be as well to mention the objections which have been urged against the use of the correction for any purpose. Karl Pearson (6) considered that the method was sufficiently condemned by the fact that it sometimes led to the absurd result of giving a value of r which was greater than one. This, as Spearman has pointed out, is simply due to the fact that the sampling error of a coefficient corrected for attenuation is greater than would be that of an uncorrected coefficient of the same size. It results from the fact that the argument for the formula assumes that the intercorrelations of the 'errors of measurement' are zero, whereas they may, in fact, have values positive or negative differing from zero within the limits imposed by their own probable errors. It would certainly be more pleasant to have a method for correcting for attenuation which had unity as its theoretical maximum for the corrected coefficient, and it is possible that such a method may be devised some time. Until it is, the possibility of getting a corrected r which is greater than unity is not a very serious defect of the method if it is remembered that the estimate of r given by the method is (like every other statistic) limited in its accuracy by its sampling error. If the corrected coefficient is correctly calculated from appropriate data, it will not be found to be significantly greater than unity.

While this is a sound (and, I think, adequate) defence against the charge that the occurrence of values of the correlation coefficient greater than unity is a fatal objection to the use of correction for attenuation, I have, nevertheless, the impression (which I cannot support by exact figures) that coefficients corrected for attenuation come out as greater than unity rather more often than they should. Since investigators frequently give corrected coefficients without giving the raw coefficients on which they are based or saying what formula was used for correction, it is not possible to be certain in what way this tendency to over-correction (if it exists) has arisen. If the formula with the arithmetic mean were used instead of that with the geometric mean, this would cause a slight over-correction, since the geometric mean of a number of unequal figures is always less than their arithmetic mean. The effect would be small unless the correlations of the numerator were very unequal. A more likely source of over-correction is the use of $r_{A_1B_1}$ and $r_{A_2B_2}$ under the conditions already mentioned in which function fluctuation makes their use invalid.

The use of correction for attenuation has also been objected to not on mathematical but on psychological grounds. It is sometimes said that the inconsistencies between two successive administrations of the same test to a group are not 'errors' but are all psychologically caused; a subject who was attentive when answering an item in one administration and got it right may make a mistake in the same item when the test is repeated because he is distracted, and so on. Variations in attention are real psychological events and not accidents. This objection has, however, no relevance to the question of the legitimacy of correction for attenuation, since this method does not assume that the so-called 'errors' are accidental. All that is necessary for the justification of the method is that the same test twice applied to the same group gives two sets of results that are not exactly the same. The individual deviations can then be regarded as the sum of two parts, one part showing perfect correlation between the two applications of the test and the other showing no correlation. The uncorrelated elements in the deviations need not be accidental; they are treated as 'error' because they are irrelevant to the purpose in hand, which is the assessment of the relative abilities of the tested individuals in whatever capacity the test is designed to measure. For other purposes, these fluctuations may be relevant; their irrelevance (the character in virtue of which they are treated as 'error') is not absolute; it is only relative to the purpose of the test.

In addition to objections to the attenuation correction in principle, there are certain particular difficulties in its application which depend on the fact that particular test results may not conform to the conditions that are assumed in the theory of the method. One such difficulty has already been discussed, that caused by day-to-day fluctuation of the function measured.

Another difficulty arises from the fact that if $r_{X_1X_2}$ and $r_{Y_1Y_2}$ are obtained in each case by repetition of the same test, it is possible that the correlation between them may be increased by the subject remembering his previous answers and repeating previous successes and failures in this way. This factor will lead to underestimation of the corrected coefficient. Since there is no reason for using correction for attenuation unless an estimate of the absolute value of the correlation coefficient is required, under-correction can no more be tolerated than over-correction. This memory effect will become less as time between the two testings is increased and may become negligible if sufficient time is allowed to elapse, partly because the subject's memory of previous responses rapidly fades, and partly because (in the case of children) increase of chronological age will make the critical questions fall in a new part of the test on the retesting. To some kinds of test, the objection does not apply at all. It is, for example, very improbable that memory would cause any intercorrelation between errors in Seashore's pitch-discrimination test even if the repetition were given immediately after the first application.

This consideration leads us to prefer for many tests the alternative of dividing it into two equivalent halves or of giving another test in the same form but with different material. If either of these methods is used, it is obviously necessary that equivalence should be as close as possible, while similarity must not be so close that the first objection applies. If the two similar tests do not resemble each other sufficiently closely, there may be expected to be a considerable specific factor in each which will lower the reliability coefficients and so enhance the value of the coefficient of correlation corrected for attenuation. Since in every case where a reliability coefficient has been calculated from similar but not identical tests it is to be expected that there will be at least a small specific factor, we may notice that this is not a serious objection to the use of such material for correction for attenuation. Let us suppose that the deviations of each test result are made up of a factor common to the other test of the same kind, of a specific factor, and of an error of measurement. Thus the deviations of X_1, X_2, Y_1 and Y_2 can be represented as $(x+s_1+e_1)$, $(x+s_2+e_2)$, $(y+s_3+f_1)$ and $(y+s_4+f_2)$, respectively. Now if the corrected coefficient had been obtained by merely repeating X_1 and Y_1, its value would have been that of the correlation between $(x+s_1)$ and $(y+s_3)$. If, however, the corrected coefficient is calculated from the two pairs of non-identical tests X_1 and X_2 and Y_1 and Y_2, the value obtained is that of the correlation between x and y. This will be larger than the value obtained by the other method, but it will be in no sense a spurious enhancement of the correction. This value also has a definite and legitimate meaning.

There is, however, a more serious objection to the use of the self-correlation of two halves of the test as a measure of the test reliability and as a means of correcting for attenuation, which applies to most (though not to all) modern group tests of intelligence. These are very generally constructed so that either the test as a whole or each section of it has a time limit such that most testees do not have time to complete all the test items. This method of construction has obvious practical advantages, but it has the unfortunate effect that the two equivalent halves into which the test items can be divided do not correspond to two independent applications of equivalent tests.

Let us suppose, for example, that a particular section of a test contains twenty items and that a particular subject manages to do twelve of them in the time limit allowed. On a retesting, he might do rather more or rather less than this, and such random variations of speed are part of what makes the self-correlation of the test fall below unity and attenuates its correlations with all other tests. It does not, however, reduce the correlation between the two halves of the test formed in such a way as that of taking the odd and the even test items, for six of each of these have been done in the time allowed for the section. If the number of test items completed were the only factor in determining individual differences in score, the 'reliability' as measured by the intercorrelation of equivalent halves would be approximately unity. Actually, scores are also determined by the number of mistakes and these will be distributed over the two equivalent halves much as they would be between two independent tests, and will cause the half-half correlation to fall short of unity, but to an extent which may considerably over-estimate the real reliability of the test.

The difficulty here is that the self-correlation of independent testings is determined by the variability of two factors—speed and correctness. Of these, correctness alone determines the intercorrelation of equivalent halves. Therefore the reliability is over-estimated to an extent which depends on the relative importance of speed and of correctness as determinants of individual differences in score. Any correction for attenuation obtained from such a 'reliability' coefficient will suffer from serious under-correction. The equivalent-halves method cannot validly be used for determining the reliability of a test which has a time limit either for the whole test or for its sections. The originators of tests of this kind have sometimes published 'reliability' coefficients obtained in this way. It is apparent both in theory and in practice that these estimates of reliability are dangerously over-optimistic.

The fact that each subject has practised the first test by the time he does the second may also cause a difference between the two sets of responses which, if the effects of the practice are different in different subjects, will lower the correlation between X_1 and X_2 and between Y_1 and Y_2 and so lead to over-correction for attenuation. Kelley (4) has suggested that this difficulty may be got over by giving sufficient fore-exercise for each test to establish the appropriate attitude and to give familiarity with its structural features.

Thus it may be seen that there are certain conditions to be observed in the collection of data for correction for attenuation. With the exercise of sufficient care, data for most tests can be made appropriate for this operation. Sometimes, probably, the difficulties are insuperable. I know, for example, of no performance scale of intelligence that can be repeated without memory largely influencing the retest result, or that can be split into two equivalent parts, or that can be considered effectively equivalent to any other performance scale. Valid correction for attenuation would be very difficult with such material. More generally the difficulties may be overcome by care in planning the design of the experiment.

III. CONDITIONS IN WHICH COEFFICIENTS OF CORRELATION SHOULD AND SHOULD NOT BE CORRECTED FOR ATTENUATION

If we have agreed that correction for attenuation is a legitimate operation when applied to correlation coefficients that have been obtained under appropriate conditions, it does not follow that this is an operation which it is always either necessary or desirable to apply in psychological research. Whether correlation coefficients should or should not be corrected for attenuation depends on the purpose for which they are required.

Let us consider then the purposes for which we may be calculating a coefficient of correlation:

(1) We may want to know how far success in one test is predictive of success in another test. It is, for example, a common problem in educational psychology to discover how successfully the results of school entrance examinations predict subsequent school success. Here the answer to our problem is given by the correlation coefficient uncorrected for attenuation and the corrected coefficient would be definitely wrong. We want to know what is the predictive value of the actual entrance examination, not what would be the predictive value of an ideal entrance examination with perfect self-consistency. A prophet must be judged by his success in foretelling the future under the actual conditions in which he has to prophesy, not by an estimate of how well he would foretell the future if these conditions were improved. Lack of self-consistency is one of the conditions under which the entrance examination makes its prediction of future success, and to eliminate the effect of this condition would be to give a falsely favourable estimate of its predictive success.

This may be stated more generally. Whenever we want to determine the amount of correlation between test measurements (and not the amount of correlation between the abilities or other mental functions measured by the tests), then the attenuated coefficient must be used.

(2) It very commonly happens that we work out a correlation coefficient in order to determine whether or not two measurable mental characteristics are causally connected. Let us suppose that we want to test the hypothesis that skills depending on hand-eye co-ordination are easier to acquire if the eyes are wider apart. One way of solving this problem would be to measure the distance between the eyes of a large group of persons and to test their skill in a number of such performances as throwing darts, catching balls, etc., and to work out the correlation coefficient between eye distance and each of these skills. Causal connexion of some kind or other is indicated by the discovery of a significant correlation, and if our purpose is only to discover which (if any) of these skills are causally dependent on eye distance, we need only to determine which of the correlations are significant. The fact of causal connexion will be made no clearer by knowing how great is the correlation between eye distance and the capacity in question. If we are interested only in the fact of causal connexion, there is no possible advantage in working out

the coefficient as corrected for attenuation; if the corrected coefficient is significant, then so also will be the uncorrected coefficient. The attenuated coefficient is, therefore, to be preferred, since it requires much less labour to determine and gives the required information equally well.

Generally, if our object is to demonstrate the existence of correlation between A and B and not to determine its amount, then the attenuated coefficient serves our purpose as well as the coefficient corrected for attenuation. The uncorrected coefficient is, therefore, the better one to use, since the additional labour of correcting for attenuation serves no useful purpose.

(3) It also frequently happens that we require a correlation coefficient in order to perform some operation with it whose validity does not depend on correction for attenuation. We may, for example, wish to discover whether a group factor of 'mathematical ability' underlies the ability in different mathematical subjects. We test a group of school children in geometry and algebra and try to determine by means of Spearman's 'tetrad equation' whether the correlation between these two abilities is greater than would be due to the dependence of both on general intelligence or G. We obtain the group scores in two tests which have no other common factor than G (let us say a verbal and a perceptual test of intelligence). Then if the tetrad difference $r_{GA} . r_{VP} - r_{GV} . r_{AP}$ is significantly greater than zero, we can conclude that ability in geometry and algebra depend on a common factor which is not G. Now Spearman has shown that a tetrad relationship which holds between coefficients corrected for attenuation will hold also between coefficients uncorrected for attenuation. Here again, therefore, it is unnecessary to correct for attenuation, and in problems of this type it is better to save labour by not doing so.

To put this in general terms. If we wish to use a correlation coefficient to demonstrate a relationship which holds equally between corrected and uncorrected coefficients, correction for attenuation is unnecessary and should, therefore, not be used.

(4) The cases so far discussed cover the greater number of the actual problems for which coefficients of correlation are required in psychological research work. There remains the case in which we wish to know the degree of correlation, not between the test results but between the abilities measured by the tests. Sometimes the problem is of the correlation between character traits or between characteristics of perception which are not of the nature of abilities. I shall, therefore, use the general term 'mental function' for what is measured by tests, meaning to include

in this term such things as abilities, character traits, perceptual characteristics (such as phenomenal regression) and so on. Where the problem is of the absolute amount of correlation between mental functions, the correction for attenuation must be used. If any psychologist feels so strongly the objections to correction for attenuation which have already been discussed that he is unable to use this correction, then he cannot use correlation coefficients for the solution of the present problem. The use of the coefficient uncorrected for attenuation as a measure of the amount of correlation between the mental functions measured by two tests is wrong and wholly unjustifiable. All that the uncorrected coefficient tells us is that the required coefficient of correlation lies somewhere between the value indicated and unity.

We might, for example, want to know to what extent ability to pass school entrance examinations depended on the same individual characteristics of children as did later school success. This would be a wholly different problem from that discussed in (1), although it would start from the same primary data. For answering this question, the coefficient of correlation must be corrected for attenuation. The uncorrected coefficient would under-estimate to an unknown extent the amount of what was common to the ability to pass the entrance examination and the ability to succeed later at school.

We certainly sometimes want to obtain information of this kind. Probably we want it less often than the psychological investigator is inclined to imagine, since there is a tendency to suppose that the coefficient so obtained gives more information than it actually does. The correlation coefficient is often spoken of as if it gave a measure of the strength of the causal connexion between two variates. In reality the information it gives is of a less fundamental character than this. To quote Fisher (8): "It [the correlation coefficient] tells us the relative importance of the factors which act alike upon the [variates in question] compared to the totality of factors at work." In other words, the correlation coefficient can be represented as the fraction $X/(X+Y)$, in which X represents the variance of the factor which is common to both tests and Y represents that part of the total variance which is specific to each test.

The magnitude of the correlation will thus depend on the relative importance of the factors specific to each test in the particular population from which the sample investigated has been drawn (i.e. on the size of Y). Inequalities of teaching and of school attendance will, for example, probably be some of the factors causing individual differences in later school success which are uncorrelated with ability to pass the school entrance examination. If we measure the correlation in a sample drawn from a population with great inequalities either of teaching or of attendance, then the correlation coefficient will be correspondingly smaller. If these inequalities are less the correlation will be greater. If these are the only factors other than initial scholastic ability which influence school success, then if there are no inequalities in them the correlation will be unity. In all cases the correlation coefficient is a measure not of the absolute importance of initial scholastic ability in determining later school success but of its importance relative to various other factors. If, as certainly sometimes happens, it is important for us to have this information, then the correlation coefficient corrected for attenuation is what we require.

We may thus see that in the practical psychological investigations for which we require correlation coefficients, it may happen (although not, I think, very commonly) that a coefficient corrected for attenuation is the only one that will give us a correct answer to our question. In the majority of our inquiries, a corrected or uncorrected coefficient will do equally well. Sometimes (although again not very commonly) an uncorrected coefficient is the only one that answers our question correctly. These facts (and especially the last) are obscured by a common habit of referring to the coefficient corrected for attenuation as the 'true correlation', with the obvious implication that otherwise the coefficients are untrue. The previous argument shows that the coefficient uncorrected for attenuation may be as true as or truer than the corrected coefficient. It is always a perfectly true measure of the degree of correlation between test measurements; it only becomes untrue when it is taken also to be a measure of the degree of correlation between the functions measured by the tests. I suggest that this can be best expressed by abandoning the term 'true' correlation for the coefficient corrected for attenuation, and substituting the term 'function' correlation. The coefficient uncorrected for attenuation can similarly be called the 'measurements' coefficient of correlation. It is not merely an incorrect estimate of the correlation between the functions measured by the two tests; it is a correct estimate (i.e. correct within the limits of its own sampling error) of the correlation between the measurements given by the two tests. Both correlations are, for different purposes, equally 'true'.

There is one other possibility which, I think, deserves consideration. Let us suppose that we wish to find out how well a newly devised intelligence test (A) predicts the general ability measured by some established test of intelligence (B), let us say, the *Stanford Revision*. Here we have a problem intermediate between that for which the 'function' correlation and the 'measurements' correlation is required. The coefficient of correlation corrected for attenuation would over-estimate the predictive value of the test, since it would tell us not how well the test does in fact diagnose intelligence as measured by the standard test but how well it would diagnose it if our new test were free from unreliability (i.e. under conditions better than those in which the diagnosis must in fact be made). If entirely uncorrected for attenuation, however, the coefficient would under-estimate the diagnostic value of the test, since it would be including, as an imperfection of the test, the unreliability of the standard. Here what we want to know is how well the measurement of our new test correlates with the function measured by the standard test; the effect of errors of measurement in the standard must be eliminated but not the effect of those in the test which is under investigation. By the same kind of argument as is used for the full correction for attenuation, it can be shown that the correlation here required is given by the formula

$$\sqrt{\frac{r_{AB_1} \cdot r_{AB_2}}{r_{B_1B_2}}}.$$

I suggest that this should be called the 'measurement-function' coefficient of correlation. Its general use is to determine how well the measurements in one test diagnose the function measured (validly but not with perfect reliability) by some other test.

That the use of a correlation coefficient (or any other statistic) should be directed towards the obtaining of an answer to a particular question may seem a point too obvious to be worth emphasizing. It is, however, surprisingly often ignored. The research worker in educational psychology too often accumulates a series of arrays of test results, and works out correlations between them, sometimes correcting for attenuation and partialling out any variates for which he has the necessary test data, all with no clearer object than the pious hope that effort in a procedure at once so respectable and so laborious must lead to some kind of useful result. It is not surprising that this hope is very commonly disappointed.

IV. PARTIAL CORRELATION

Another correction commonly applied to correlation coefficients is for eliminating the effect of the partial correlation of either or both variates with some other variate. We have, let us say, given tests in two school subjects to a group of children from different forms in a school. Those from higher forms will, on the whole, do better in both tests because they have studied the subjects longer and at a higher standard than the others. This fact will make the correlation higher than it would have been if all the children had been of the same school standing. We may wish to know what the correlation between the tests would have been if all the children had been of the same school standing. This is determined by the use of Udny Yule's partial correlation formula:

$$r_{AB(C)}=\frac{r_{AB}-r_{AC}\cdot r_{BC}}{\sqrt{(1-r_{AC}^2)\,(1-r_{BC}^2)}},$$

in which r_{AC} and r_{BC} are the correlations of the two tests with school standing and $r_{AB(C)}$ is the required correlation between the two tests if school standing is held constant (that is, in a population homogeneous with respect to school standing). Generally the corrected coefficient will be less than it was before if the eliminated variate affects both A and B in the same direction, but the correction will increase the coefficient if the eliminated variate affects one of the correlated variates but not the other, or if it affects both in opposite directions. In such an inquiry as this it may, of course, happen that we want not only to eliminate the effect of differences in school standing but also in chronological age, or in intelligence or in both.

Let us suppose that we wish to eliminate the effect on r_{AB} of differences in intelligence. We are then faced by a very serious limitation to the use of this formula which is not always realized by psychologists who want to work out a partial correlation; the formula is only valid if the eliminated variable C can be measured without error. No test result fulfils this condition and, therefore, this formula can never be used to partial out any test result unless all the correlations used have been corrected for attenuation. It seems that the only two things we can measure without error in educational psychology are chronological age and school standing. Neither of these will, it is true, be normally distributed in any results we are likely to get in a practical investigation, but M. S. Bartlett (7) has shown that normal distribution is not necessary in the eliminated variate in a partial correlation. The partial correlation coefficient can, therefore, be used as it stands for eliminating the effect of chronological age and of school standing if these are irrelevant factors for the purpose for which we require the coefficient. Otherwise its use is illegitimate and will lead to misleading results.

When we say that the use of the ordinary partial correlation formula is misleading when the measurements of the eliminated variate are subject to error, what is meant is not merely that its use gives an incorrect estimate of $r_{AB(C)}$. That is obvious, and, for many purposes, may be unimportant. A much more serious objection to the use of this formula in this case is that it cannot be relied upon to provide accurate information as to whether A and B are correlated at all in a population homogeneous with respect to C. That, at least, we want to know when we work out a partial correlation coefficient. As will be shown later, however, the ordinary formula may (and normally will) give us a value of $r_{AB(C)}$ differing from zero if C is subject to errors of measurement when, if C had been measured without these errors, $r_{AB(C)}$ would have been zero. The ordinary formula cannot, therefore, in this case, even be relied on to tell us whether A and B are correlated when C is held constant. In other words, if C is appreciably affected by errors of measurement, the application of the ordinary formula without any correction for these errors of measurement is of no use whatever.

An obvious way out of this difficulty is to obtain duplicate testings of A, of B and of C, and to correct all the coefficients for attenuation by use of Spearman's formula. This is legitimate but laborious. Occasionally it is not possible, since some tests (e.g. some of the performance scales of intelligence) cannot be repeated and cannot be split up into equivalent halves. If this procedure of complete correction for attenuation is carried out, it leads to the discovery of the 'function' correlation between A and B when C is constant. This is generally satisfactory, but it may sometimes happen that it is the 'measurements' correlation between A and B that is required.

I suggest that when the 'measurements' correlation is needed or when it will serve our purpose just as well as the 'function' coefficient, a method of dealing with the problem of partial correlation less radical than that of complete correction for attenuation should be adopted. The self-correlation of C must be known, so there must be double measurement of C but not of A or B. The required partial correlation between A and B is then given by the formula

$$\frac{r_{AB} \cdot r_{C_1C_2} - \sqrt{(r_{AC_1} \cdot r_{AC_2} \cdot r_{BC_1} \cdot r_{BC_2})}}{\sqrt{\{(r_{C_1C_2} - r_{AC_1} \cdot r_{AC_2})\,(r_{C_1C_2} - r_{BC_1} \cdot r_{BC_2})\}}}.$$

The proof of this formula is given in the Appendix.

The relationship of this formula to the usual formula for partial correlation can be made most clear if we remember that r_{AC_1} will be approximately the same as r_{AC_2} and that r_{BC_1} will be approximately the same as

r_{BC_2}. So $\sqrt{(r_{AC_1} r_{AC_2})}$ and $\sqrt{(r_{BC_1} r_{BC_2})}$ may be written as r_{AC} and r_{BC} respectively. The formula then becomes

$$\frac{r_{AB}\cdot r_{C_1C_2} - r_{AC}\cdot r_{BC}}{\sqrt{\{(r_{C_1C_2}-r_{AC}^2)\,(r_{C_1C_2}-r_{BC}^2)\}}},$$

from which it is obvious that if $r_{C_1C_2}$ were unity (i.e. if there were no error of measurement in test C), the formula would reduce to the usual form for partial correlation.

The purpose of the formula we are at present discussing is to give the partial correlation between the *measurements* of A and B in a population homogeneous with respect to C. It is suggested that for purpose (1) (of the earlier discussion, on p. 392), it is the only right partial correlation formula to use if C has errors of measurement, that for purposes (2) and (3) it is preferable to the method of complete correction for attenuation since it is less laborious, while for purpose (4) it is inadmissible and the method of complete correction for attenuation must be used.

It might be supposed that since it has been argued that the total correlation coefficient uncorrected for attenuation has a definite meaning (that of correlation between test results) and is the proper measure of correlation for some purposes, the partial correlation coefficient calculated from the usual formula when the eliminated variable is subject to uncorrected errors of measurement might be justified for certain purposes by a similar argument. Has not this coefficient also a definite statistical meaning? Certainly it has, but it is not a meaning that is of any practical use to us. It is a measure of the correlation that would exist between A and B in a population in which the test scores in test C were constant. I do not think that there is any conceivable inquiry in which this would be a useful piece of information. Perhaps the greatest use of the attenuated coefficient of correlation is the second use mentioned above—to tell us whether the two tests in question are correlated. If the attenuated correlation coefficient is significant, then the functions measured by the two tests are correlated. Correction for attenuation will give us a better estimate of how much they are correlated (if we require to know this) but it will give us no more certain indication of the fact of correlation.

There is no similar justification for the use of the uncorrected partial correlation coefficient. Let us suppose that two abilities A and B are correlated in a population heterogeneous with respect to C, and that our reason for using the partial correlation method is to discover whether or not they are correlated in a population homogeneous with respect to C. We cannot infer this from observing whether or not they are correlated in a population homogeneous with respect to C scores, since they may be correlated in a population homogeneous with respect to C scores although uncorrelated in a population homogeneous with respect to the function C itself. This is clear if we consider the case that A and B are uncorrelated in a population with C constant but are correlated in a population heterogeneous with respect to C. Then, by the formula last given, we know that $r_{AB}.r_{C_1C_2}-r_{AC}.r_{BC}=0$. Since, however, $r_{C_1C_2}$ is less than one (if the test of C is not perfectly self-correlated), it follows that $r_{AB}-r_{AC}.r_{BC}$ must be greater than zero. It follows that the usual formula will make A and B appear to be correlated when C is held constant although, in fact, they are uncorrelated. Obviously it is also theoretically possible (although less likely to happen in practice) that this formula will make A and B appear to be uncorrelated when they really are correlated.

The usual formula is, therefore, to be condemned as a method of obtaining partial correlations whenever the variate partialled out is subject to appreciable errors of measurement. This means that it is to be condemned for most purposes for which it is likely to be required in psychology. It is necessary, either that A, B and C should be measured twice and all coefficients corrected for attenuation, or that C alone should be measured twice and that the formula I have suggested should be used. The first procedure will give the 'function correlation' between A and B, and the second will give the 'measurements correlation' between A and B, both with C held constant. Which procedure will be adopted must depend on the purpose for which the coefficient is required, as has been discussed earlier.

V. SUMMARY

Correction for attenuation is a justifiable operation when properly applied to appropriate data, although it is not always (or most commonly) necessary.

Certain precautions are necessary when applying it to coefficients of correlation between mental functions which show day-to-day fluctuation.

The required estimate of reliability may be obtained: (A) by repeating the same test, (B) by splitting it into two equivalent halves, or, (C) by repeating a test of equivalent difficulty. (A) is not appropriate if similarity of responses in test and retest might be caused by memory. (B) is not appropriate if the test has a time limit.

Correction for attenuation is wrong if what is required is the correlation between two test scores. It is not necessary if we require only to know that two tested functions are correlated, or if we want to use the coefficients for the investigation of some relationship which holds equally between attenuated and unattenuated coefficients. It is necessary only when we want to know the absolute value of the correlation coefficient between the two tested functions. Consideration of the meaning of the correlation coefficient suggests that this knowledge is probably less often of importance than is commonly supposed.

The usual formula for correction for partial correlation is not valid if the eliminated variate is subject to errors of measurement. This severely limits its usefulness in psychological investigation. A formula is suggested for the valid application of the correction for partial correlation when it is not desired (or not possible) to correct all the coefficients used for attenuation.

APPENDIX

Proof of the formula for obtaining the 'measurements' correlation between A and B, with C held constant, C being itself subject to errors of measurement.

Four test measurements must be made: A, B, C_1 and C_2 (i.e. test and retest of C).

Let a and b be the elements in the deviations specific to A and B respectively (including their errors of measurement).

Let n be the element in the deviations of A and B which is common to them but not to C.

Let m be the element in the deviations of C which is found also in A and B with weights represented by pm and qm, respectively.

Let c_1 and c_2 be the elements in the deviations of C_1 and C_2, respectively, which are not common to both measurements of C (i.e. their respective errors of measurement).

Then the deviations of A are compounded of $a+pm+n$.
" " " B " " $b+qm+n$.
" " " C_1 " " c_1+m.
" " " C_2 " " c_2+m.

The partial correlation required is the correlation between A and B with m held constant, that is, $r_{(a+n)(b+n)}$.

Now $$r_{(a+n)(b+n)}\cdot\sigma_{(a+n)}\cdot\sigma_{(b+n)}=r_{ab}\sigma_a\sigma_b+r_{an}\sigma_a\sigma_n+r_{bn}\sigma_b\sigma_n+r_{nn}\sigma_n^2$$
$$=\sigma_n^2$$

(since, by hypothesis, a, b and n are uncorrelated, while $r_{nn}=1$).

So $$r_{AB(m)}=\sigma_n^2/[\sigma_{(a+n)}\cdot\sigma_{(b+n)}]. \qquad \ldots\ldots(1)$$

Similarly, $$r_{AC_1}\sigma_A\cdot\sigma_{C_1}=p\cdot r_{mm}\sigma_m^2=p\cdot\sigma_m^2$$

and $$r_{AC_2}\sigma_A\cdot\sigma_{C_2}=p\cdot\sigma_m^2,$$

so $$r_{AC_1}r_{AC_2}\sigma_A^2\cdot\sigma_{C_1}\cdot\sigma_{C_2}=p^2\cdot\sigma_m^4; \qquad \ldots\ldots(2)$$

similarly, $$r_{BC_1}r_{BC_2}\sigma_B^2\cdot\sigma_{C_1}\cdot\sigma_{C_2}=q^2\cdot\sigma_m^4. \qquad \ldots\ldots(3)$$

Also
$$r_{C_1C_2}\sigma_{C_1}\sigma_{C_2}=r_{c_1c_2}\sigma_{c_1}\sigma_{c_2}+r_{c_1m}\sigma_{c_1}\sigma_m+r_{c_2m}\sigma_{c_2}\sigma_m+r_{mm}\,\sigma_m^2$$
$$=\sigma_m^2 \qquad \text{......(4)}$$

(since, by hypothesis, c_1, c_2 and m are uncorrelated).

In the same way, we find that

$$r_{AB}\,\sigma_A\,\sigma_B=\sigma_n^2+pq\cdot\sigma_m^2, \qquad \text{......(5)}$$

so
$$\sigma_n^2=r_{AB}\,\sigma_A\,\sigma_B-pq\cdot\sigma_m^2.$$

Substituting this in (1), we get

$$r_{AB(m)}=(r_{AB}\,\sigma_A\,\sigma_B-pq\cdot\sigma_m^2)/[\sigma_{(a+n)}\,\sigma_{(b+n)}]. \qquad \text{......(6)}$$

Multiplying (2) and (3) and taking the square root:

$$pq\cdot\sigma_m^4=\sqrt{(r_{AC_1}\cdot r_{AC_2}\cdot r_{BC_1}\cdot r_{BC_2})}\cdot\sigma_A\,\sigma_B\,\sigma_{C_1}\,\sigma_{C_2}.$$

Dividing by (4), this becomes

$$pq\cdot\sigma_m^2=\sqrt{(r_{AC_1}\cdot r_{AC_2}\cdot r_{BC_1}\cdot r_{BC_2})}\cdot\sigma_A\,\sigma_B/r_{C_1C_2}. \qquad \text{......(7)}$$

By the ordinary formula for summing independent variances

$$\sigma_A^2=\sigma_{(a+n)}^2+p^2\sigma_m^2.$$

So
$$\sigma_{(a+n)}^2=\sigma_A^2-p^2\sigma_m^2 \text{ and } \sigma_{(b+n)}^2=\sigma_B^2-q^2\sigma_m^2.$$

Therefore
$$\sigma_{(a+n)}\cdot\sigma_{(b+n)}=\sqrt{\{(\sigma_A^2-p^2\sigma_m^2)\,(\sigma_B^2-q^2\sigma_m^2)\}}. \qquad \text{......(8)}$$

If we divide each side of (2) and of (3) by (4), we get

$$p^2\sigma_m^2=r_{AC_1}\cdot r_{AC_2}\cdot\sigma_A^2/r_{C_1C_2} \text{ and } q^2\sigma_m^2=r_{BC_1}\cdot r_{BC_2}\cdot\sigma_B^2/r_{C_1C_2}.$$

Substituting these values in (8):

$$\sigma_{(a+n)}\cdot\sigma_{(b+n)}=\sigma_A\,\sigma_B\sqrt{\{(1-r_{AC_1}r_{AC_2}/r_{C_1C_2})\,(1-r_{BC_1}r_{BC_2}/r_{C_1C_2})\}}. \qquad \text{......(9)}$$

Substituting (7) and (9) in (6), and multiplying numerator and denominator by $r_{C_1C_2}/\sigma_A\cdot\sigma_B$:

$$r_{AB(m)}=\frac{r_{AB}\cdot r_{C_1C_2}-\sqrt{(r_{AC_1}\cdot r_{AC_2}\cdot r_{BC_1}\cdot r_{BC_2})}}{\sqrt{\{(r_{C_1C_2}-r_{AC_1}r_{AC_2})\,(r_{C_1C_2}-r_{BC_1}r_{BC_2})\}}}.$$

This is the formula for the partial correlation between the measurements of A and B when the variate partialled out is subject to errors of measurement.

If there were no errors of measurement in C, $r_{C_1C_2}$ would be unity, and

$$r_{AC_1}=r_{AC_2}=r_{AC} \text{ and } r_{BC_1}=r_{BC_2}=r_{BC},$$

so the above formula would then be reduced to the usual form:

$$r_{AB(C)}=\frac{r_{AB}-r_{AC}r_{BC}}{\sqrt{\{(1-r_{AC}^2)\,(1-r_{BC}^2)\}}}.$$

REFERENCES

(1) SPEARMAN, C. (1904). "'General Intelligence' objectively determined and measured." *Amer. J. Psychol.* XV, 201.
(2) —— (1907). "Demonstration of formulae for true measurement of correlation." *Amer. J. Psychol.* XVIII, 161.
(3) —— (1910). "Correlation calculated from faulty data." *Brit. J. Psychol.* III, 271.
(4) KELLEY, T. L. (1923). *Statistical Method.* New York.
(5) THOULESS, R. H. (1936). "Test unreliability and function fluctuation." *Brit. J. Psychol.* XXVI, 325.
(6) PEARSON, K. (1904). "On the laws of inheritance in man" (addendum on p. 160). *Biometrika*, III, 131.
(7) BARTLETT, M. S. (1932). "On the theory of statistical regression." *Proc. roy. Soc. Edinb.* LIII, 260.
(8) FISHER, R. A. (1925). *Statistical Methods for Research Workers.* Edinburgh.

(*Manuscript received* 24 *May* 1938)

THE GENERAL FACTOR IN CORRELATIONS BETWEEN PERSONS

By M. DAVIES

From the Psychological Laboratory, University College

I. *Problem* (pp. 404–405).
II. *Methods* (pp. 405–409).
III. *Data* (pp. 409–412).
IV. *Results* (pp. 412–416).
V. *Summary of conclusions* (pp. 417–420).
  A. *As regards methods* (p. 417).
  B. *As regards results* (pp. 417–420).
References (p. 420).

I. PROBLEM

It is a familiar fact that, on correlating mental tests, the majority of the correlations prove to be positive, and that a single general factor will account for by far the larger proportion of the variance. It is therefore natural to enquire whether the same remarkable result will reappear when we correlate, not tests, but persons. To this question two strongly conflicting answers have been given.

Most earlier investigators have apparently assumed that the situation should be much the same in both cases. This view has been explicitly maintained by Burt (20) and Dewar (25); and is implied by the work of Bulley (see (9)), Beebe-Center (10), Pritchard (12), Shakespeare (19), and many others. Stephenson, on the other hand, in a recent series of papers (15, 16, 17) has vigorously contested this point of view. He insists upon a sharp opposition between the two methods of approach. Where correlations between tests are concerned, he still upholds the doctrine of a single general factor; but, where correlations between persons have been calculated, he finds, in nearly every case, not one general factor but two group factors, each limited to a particular group of persons, and together producing a pair of mutually exclusive 'types'.

If this conclusion could be established, several important consequences would follow. For example, the indiscriminate averaging of personal orders or rankings (the method adopted in so many studies of aesthetic preferences and the like) would become an "illegitimate and meaningless procedure" ((14), p. 44). Moreover, the conclusion itself would seem, at least at first sight, to be quite incompatible with the main tenet of the Spearman school. On this ground, indeed, Stephenson and others originally rejected the method of correlating persons. He now, however, maintains that the factors revealed by the two methods "must belong to entirely different worlds", and supports his belief by a number of highly suggestive arguments. Nevertheless, before we decide to accept either of these rather drastic corollaries, the conclusion itself deserves a more searching scrutiny.

Curiously enough, no one has hitherto undertaken a comprehensive survey of the existing data in the way Spearman once surveyed all the available tables obtained by correlating tests. The object of this paper is to present the results of such a review, in the hope of reconciling antagonistic opinions on this and several minor but cognate controversies.[1]

II. METHODS

To a large extent the differences of view seem attributable merely to differences in experimental or statistical procedure. In this country, for example, most investigators who have factorized correlations between persons have followed one or other of the multiple-factor methods originally outlined by Burt for that purpose (in his laboratory *Notes on Factor Analysis*, afterwards partly incorporated in (13), Memorandum I); and have used either his method of least squares (when dealing with covariances) or his summation formula (when dealing with correlations), seeking as close a fit as possible by means of successive approximation.

---

[1] It will be evident that I am deeply indebted to very many persons for their assistance. More particularly, I must acknowledge my gratitude to Dr Stephenson, formerly Research Assistant at University College. It was originally planned that my own investigation should be carried out under his immediate supervision, but his later duties at Oxford prevented any full or personal discussion; nevertheless, I am greatly in his debt for the trouble he has taken in sending detailed and stimulating criticisms of my results and for the kindly way he has always received incidental enquiries or criticisms of my own. He has also allowed me to make use of his own unpublished notes on Q-technique and of recent unpublished tables compiled by others working under his guidance (Miss Knowles, Mr Cohen, and Miss Hill). I am further indebted to Prof. Burt for allowing me to make full use of his notes on factor analysis and correlating persons, and for procuring material collected by his former students at the Institute of Education and elsewhere; to Prof. Valentine for putting me in touch with similar researches carried out at the University of Birmingham; to Prof. Beebe-Center and his statistical assistant at Harvard, Dr Harsh, for their helpful replies to my enquiries about work in America; and finally to many other investigators who have so willingly sent me tabulated data, whenever the figures in their publications or their theses were not sufficient for my purpose.

In America Beebe-Center at first[1] employed the earlier and simpler two-factor method devised by Spearman for analysing correlations between tests, and 'inverted' it for persons. Stephenson, in his first short paper on testing artistic taste, followed Burt and Bulley as regards both problems and methods: but in his longer article in this *Journal* (15), though the detailed formulae given are identical with those of Burt for a *single* factor, the proof by 'inversion' follows that of Beebe-Center.

His incidental criticisms of these two writers indicate the essential changes in his statistical approach. First of all, whereas Burt, Dewar, and even Beebe-Center (so I gather) in his later work, believe that an analysis by a multiple-factor technique is essential if additional factors are present, Stephenson maintains that "these generalized forms of factor analysis, such as that developed by Thurstone, have an artificiality which I cannot meet" ((15), p. 350). Secondly, whereas Burt maintains that, when we are correlating persons, we are in effect simply dealing with the 'transpose' of the initial matrix and therefore must reach practically the same factors as would be obtained by correlating traits, Stephenson now insists that the two systems "are the very poles apart". Burt's method, he asserts, like those of Spearman and of other psychometrists, can logically issue only in a complete denial of psychological types and type factors ((16), p. 295).

After reading my earlier draft of this section Dr Stephenson has been good enough to send a fuller explanation of his most recent views.[2] From his articles alone it is not always clear where he is merely simplifying Burt's original analysis for the aid of the reader unfamiliar with matrix algebra, and where he is definitely departing from it. Hence some of his pronouncements seemed at first sight a little inconsistent. (E.g. in his first paper on "Inverted Factor Technique" ((15), p. 345) he apparently follows Burt in assuming that the (normalized) factor matrix obtained on correlating persons may be regarded as an inverted form of the population matrix obtained by correlating tests (i.e. in the conventional symbols $L_p = P_t^{-1}$). But his reference to 'inverting' the initial matrix—and the theorems—is surely a misunderstanding. In substituting 'P-axes' for 'T-axes', Burt speaks of 'inverting' not the initial matrix (M_3), but the population and factor matrices M_1 and M_2 ((13), p. 253; footnote 1), i.e. $LV^{\frac{1}{2}}$ and P: no doubt the original choice of symbols was partly responsible for the confusion: in both 'P-technique' and 'Q-technique' we start not by 'inverting' but merely 'transposing' M_3.) Again, in his discussion of 'experiment 4' ((15), pp. 359–60; Cohen's investigation) he fully accepts the view expressed in Cohen's thesis that Q-technique will merely enable us to "find those persons who have led to the factors in R-technique and determine how many they are". And in his later article ((18), p. 195) he still speaks in the text of the 'second system' as being the "obverse or inversion" of the first; he also accepts Burt's proposal to standardize the matrices in terms of their 'unitary standard deviation' ($\sigma\sqrt{N}$) instead of the ordinary standard deviation (σ) ((13), p. 272)—a device which renders the matrix equations much simpler. But in the postscript (p. 206) he insists that "system 2 is *not* to be regarded as the 'obverse' or mere 'transpose' of data already analysed by system 1": and the preliminary standardization, he considers, must destroy any predictable relation between the two factor and population matrices deduced by the alternative approaches. In explanation he now points out that in the body of his article he was merely concerned to distinguish the main methods of standardizing and correlating the data in any matrix. This was already in type before Burt sent him his later manuscript(20), which was written just before his own postscript. The postscript, he explains, was thus added "in reply to the new arguments which seemed to destroy the uniqueness of the very method they were describing"; and is to be taken as his own "final view on this particular issue". I gather, too, that, although in his account of 'Q-technique' he employs the matrix notation and (so far as possible) the symbols used by Burt and Thomson (in order to make the points of agreement and divergence clearer), he nevertheless has sometimes had to give the terms and symbols a special meaning of his own: thus, they must be "understood from their context rather than from a textbook of matrix algebra". He adds: "If no additional factors were brought to light by correlating persons, then person-factors can have no interest"; and the idea of deriving factors from "unstandardized scores and covariances" he rejects most emphatically as "scarcely meriting a moment's serious consideration". For this reason I have thought it necessary to base my own analyses throughout on correlations and not on covariances, even where the original investigators would have preferred the latter procedure. [In passing I must again express my indebtedness to Dr Stephenson for the fullness and readiness with which he has answered questions on these points.]

---

[1] In his latest paper ((22), p. 353) he suggests substituting Thurstone's method of factor analysis; and, from a private communication kindly sent me by his statistical assistant, Dr Harsh, I gather that, in a series of experiments, not yet published, this method has indicated at least four factors operating in judgements on aesthetic value.

[2] This paragraph forms a later insertion, summarizing subsequent correspondence between Dr Stephenson and the author.

Similar divergences, not quite so clearly emphasized, are to be noted on comparing the procedures of other investigators; and must largely account for the conflicting results to which they have been led. But in addition to differences in method there seem also to be differences in the kind of test material chosen and the kinds of persons tested. Most investigators have worked (or assumed that they were working) with random samples of material and with random samples of the population; a few, however, have at times employed highly selected samples, in some cases exceptionally heterogeneous, in others exceptionally homogeneous. Now selection of itself can either eliminate 'factors' or introduce them. Here again, therefore, we have an unsuspected explanation for the contradictory conclusions. Stephenson, for instance, working in the same laboratory, used for his preliminary trials the pictures, cards, and lantern slides collected by Burt and Dewar (originally, I believe, in conjunction with Miss Bulley): but he was quickly led (so he explains) to "discard this material as unsuitable to my purpose" (the demonstration of group factors and mental types). The material substituted (e.g. simple colours, such as those used by Bullough in his studies on artistic types ((9), p. 280)) was usually far more homogeneous; and the altered figures show, what was after all to be expected, that this change was itself tantamount to partialling out one general factor. On the other hand, his groups of persons seem much more abruptly heterogeneous than those correlated by previous investigators. One group, for instance, includes five manic-depressive patients, five schizophrenic patients, and six normal members of a single family—two elderly and four relatively young ((17), p. 363); and this experiment, as we shall see, is one of the very few in which the specific 'type factors' are statistically significant.

The prior selection of types—of groups within groups, as it were—raises yet another issue. In analysing correlations between persons, as in analysing correlations between tests, the investigator has always to decide between two main alternative procedures—between what may be called the 'group-factor method' and the 'general-factor method' of analysis respectively ((13), p. 306). (*a*) With the first method, we partition the entire matrix of correlations into submatrices, and seek a separate factor for every submatrix that contains positive coefficients higher than those of the adjacent submatrices; this yields two or more distinct factors, each sharply confined to a certain group of persons, and possessing for those persons positive saturations only, and zero for the rest. (*b*) With the second method we analyse the matrix as a whole: this yields one or more general or 'universal' factors, shared by every person correlated. Whether tests or persons are correlated, most investigators have hitherto followed the second procedure (*b*). Burt apparently uses method (*a*) in correlating mental tests (e.g. (2), pp. 55–7) and method (*b*) in correlating persons, though in the earliest experiment on correlating persons (1), where marked visualizers, audiles and motiles were compared to demonstrate imaginal types, he apparently employed a group-factor method. Stephenson, on the other hand, invariably uses method (*b*) in deducing factors from tests, since "they are by definition universal" (i.e. general factors), but appears as a rule to adopt Burt's method (*a*) when correlating persons.

The grounds for the choice are seldom distinctly stated. Usually, it would seem, method (*a*) is adopted when there are negative coefficients in the correlation matrix. This, however, is not the true criterion. Contrary to a widespread notion, even when negative correlations appear, method (*b*) can in general still be applied; the result is then a bipolar factor. But, when we are correlating persons, the negative saturations that such a factor implies are no longer devoid of meaning, as they are generally supposed to be when correlating mental abilities. The proper criterion would seem to be rather this: method (*a*) is to be applied, when and only when there is definite evidence that the variables correlated form discontinuous groups—e.g. when the selected sample forms a highly heterogeneous population ((2), pp. 55–7, (24)). It has, however, recently been shown that, when the two alternative methods are applied to one and the same table of correlations, the two resulting factor patterns are related to one another in a very simple way, expressible by a triangular matrix; this implies that the positive and negative general factors really represent differences between each positive group factor and the remainder ("Factor analysis by submatrices", (24), p. 359).

In order that my own figures may be comparable throughout, I have used method (*b*) for every table, and method (*a*) in addition for those few cases where it seemed more especially appropriate. In calculating factor variances, I have used Burt's summation-method, inserting by successive approximation the full communalities. The results have been checked by proceeding to reiterated table-by-column multiplication, until as close a fit as possible has been attained. Since (as we shall see) most of the tables are nearly hierarchical, the variances for the first factor as thus computed here differ hardly at all from the figures that would be obtained by a simple summation method such as Thurstone's or Spearman's.[1]

---

[1] Where more than one factor is significant, the saturation coefficients thus found by the above method differ from Thurstone's chiefly in that the saturation coefficients are required to be uncorrelated. They are virtually those that would be reached by Hotelling's method or

To estimate the evidence for the *presence* of more than one factor, I have first examined the standard error of the residual correlations (or rather of their inverse hyperbolic tangents) according to Fisher's criterion for small samples ((11), p. 190). The residual is essentially a difference between two correlations—an observed and a theoretical. We therefore require the standard error of this difference. But the theoretical value is given by hypothesis, and so has no sampling error. Hence, if the traditional method is used, the residuals must be judged by the probable error of the observed correlation, not by that of the residual or of zero, as most investigators seem to have supposed. In Table I (p. 413) those investigations which yield significant residual correlations, on which a second factor can be based, are marked with an asterisk.

To estimate the *importance* of the various factors, the simplest method is to state the percentage of variance contributed by each. Although in most researches the evidence in favour of specific or group factors turns out to be 'non-significant', this does not prove (as is so often assumed) that such factors do not exist; the odds in favour of their existence nearly always remain higher than those in favour of their non-existence. The issue therefore becomes not: Is there or is there not a second factor? but: How large and how significant a part does it play as compared with the first or general factor? We may therefore agree with Stephenson in rejecting Beebe-Center's use of the tetrad-difference criterion and in seeking a 'more adequate method'. But precisely the same reasons compel us to reject the use of single-factor formulae (the so-called 'two-factor method') and to seek some means for estimating the relative importance of the multiple factors.

III. DATA

In accordance with the foregoing principles I have made a complete factor analysis of all the tables of correlations between persons to which

---

by Kelley's, if *communalities* were substituted for reliabilities or unity. Except for the determination of appropriate communalities, the calculations are very much shorter than those described by Hotelling or Kelley. In computing the 'higher moments' ((21), pp. 185 *et seq.*), 'table-by-column multiplication', i.e. $R \times (Rt)$, can be substituted for 'table-by-table multiplication', i.e. $(R \times R)\,t$, unless the figures are to be taken to many decimal places. Where variances alone are needed, one may begin with table-by-column multiplication (as required for the shortened method of 'higher moments'), and then proceed to a column-by-column multiplication; thus, throughout Table I, I have taken

$$v = t'R^3 . R^3 t / t'R^3 . Rt,$$

where t' as usual denotes the summation-operator [1, 1, ..., 1]. (The detailed working instructions are given in Burt's *Notes on Factor-Analysis*, which can be obtained from the Psychological Laboratory, University College.)

I could gain access. In all they number forty-eight. Spearman's first enquiry upon correlations between tests was carried out on fourteen (see (7), p. 138), and his second on twenty (see (7), p. 139), tables; so that it is fully time that a corresponding survey was carried out for correlations between persons. The tests surveyed by Spearman were almost exclusively tests of intellectual or cognitive abilities. Here, however, as will be seen from the headings in Table I (p. 413), the mental traits examined range over a much wider field.

The earliest work on correlations between persons was (as might be expected) concerned more with cognitive than with conative capacities. At the outset the method seems to have appealed mainly to educational psychologists interested in problems of group factors and mental types, and not wholly satisfied to ascribe all intellectual differences (except what was specific to each test) to a single factor, *g*. The earliest use of the method appears to be that of Burt and Davies in 1912 for a study of imagery types among school children ((1), p. 373); and the first table of this kind to be printed in full and examined from the standpoint of factor theory is that discussed by the former in his 1921 Report on *Mental and Scholastic Tests* and based on data obtained with tests of intelligence. An allied method was also used to obtain 'profiles' for various types of scholastic disability, based on ranking test items (or school subjects) in order of difficulty, performance, or preference, for individual children: in this earlier work, however, factor analysis, as now understood, was carried out only in exceptional cases. The majority of the earlier tables may be regarded as extended sets of consistency or 'reliability' coefficients, usually for intellectual abilities or attainments: in 1926, for example, Thomson & Bailes (6) published three tables of correlations between teachers' marks; though they do not actually work out saturation coefficients by the summation method, they give (according to a common practice at that time) correlations with the average or total order, which amounts to much the same thing. Some of the most interesting data are 'consistency coefficients' obtained from one and the same individual with one and the same set of test material. From 1922 onwards, first at the National Institute of Industrial Psychology, and later at University College, Burt and his research assistants, following a fairly systematic programme, have used the method for investigating temperamental types, aesthetic appreciation, and the like. Consequently, more recent researches, as will be seen from the list in Table I, deal rather with assessments of emotional tendencies or with subjective preferences—artistic, musical, literary, scholastic, and sensory.

The recent American work, carried out chiefly at Harvard, has also been concerned with 'preferential judgements' or 'likings', or with the reverse (i.e. 'dislikes', e.g. susceptibilities to different forms of annoyance: C. M. Harsh, unpublished communication). The 'preferential judgements' were concerned with preferences for odours (10) and, in later studies, for geometrical shapes, vases, poetry, and music (22). Similar work has been carried out by Pratt at Rutgers College and by Wilson at Western University, London (Ontario). I should add that the recently published experiments, carried out by Beebe-Center's students (22) and included in my table, were primarily designed to test Birkhoff's 'aesthetic measure' in various fields. Since the writers have explicitly refrained (p. 353) from applying any form of factor analysis (doubtless because the number of items correlated was so few), it might seem unwarrantable on my part to factorize their tables and include the factor variances in my list below. In venturing to do so, my excuse is first that the results are strikingly similar to those obtained by other investigators who have applied factor analysis for themselves, and secondly that those who have professed to discover significant factors have often been content with correlating items that were just as few in number.

Stephenson's starting-point was Burt's 1935 Memorandum on factorizing evaluations of essays, etc., along the lines worked out for literary and aesthetic appreciation generally. He accepted the device of correlating persons (so-called *P*-technique) in principle, but strongly criticized the multiple factor technique, and proposed to substitute a modified technique of his own, which, he maintains, should lead to "an entirely new branch of psychometry, for which the term Personalistics may be coined". "There is nothing *statistically* novel", he writes, "about *Q*-technique, as there is in Burt's work on Person-factors; but the *psychological* implications of these factors have never been clearly appreciated, otherwise he would not have insisted on his reciprocity principle."[1] Accordingly, he himself decided to work over the same fields—aesthetic and educational preferences, emotional and clinical types—to demonstrate that the two sets of factors (the old *P*-factors and the new *Q*-factors) "have nothing whatever in common". The '*P*-technique' as used by previous workers erred by "still thinking in terms of the abilities or mental tendencies of *R*-technique": they consequently discovered only a few general factors instead of the innumerable type factors which *Q*-technique reveals. Hitherto one of the strongest supporters of Spearman's single factor theory, he now appears as one of the keenest defenders of multiple type factors. "Almost everything we factorists have denied about type psychology has been beside the point: types exist in great numbers" ((18), p. 209). He claims it as one of the chief merits of his modification that the investigations which, with the cumbersome method of multiple analysis, take the ordinary research student a couple of years can be "performed within an hour and factorized in two more". His own published work with the new method was carried out within a single year. As he himself points out, it should in consequence be regarded as consisting of "short extracts of experiments rather than completed studies"..."intended rather to illustrate a more adequate point of view than to draw final conclusions". So far, however, he himself has not yet extended his method to the cognitive field, with which his name has so long been connected—to the problems of verbal, perceptual, and perseverative types and factors, which he has hitherto investigated by *R*-technique (correlating tests). Nevertheless, his preliminary experiments, though brief, are sufficient, he believes, to reveal that "no one has previously used the technique of correlating persons either completely or with full understanding"; Thomson (in a recent paper criticizing Stephenson's rejection of group factors when correlating tests) had, he admits, "been very near to it, but saw only difficulties instead of a flood of immediate applications" ((15), p. 352).[2]

It is plain, therefore, that one of the primary objects of this paper must be to

[1] Memorandum on '*Q*-technique' (privately communicated: here and elsewhere I am indebted to Dr Stephenson for permission to quote from his unpublished notes and communications).

[2] A fuller history of the method of correlating persons is rendered unnecessary by Vernon's thorough survey (recently published) of the various methods of assessing personal qualities ((30), esp. pp. 37–41, 63–5, 77: cf. also the historical introduction prefixed to (18)).

enquire how far the arguments and examples adduced by Stephenson do really, as he himself contends, at once "invalidate all previous efforts and overcome their difficulties". The fundamental issue that he raises might, I think, be stated as follows: When we correlate persons instead of tests, is the greater proportion of the factorizable variance still attributable to the first or 'general' factor (i.e. to one g and n specifics), or is it essential to invoke further factors in addition, each contributing a large enough proportion of the variance to be statistically significant?

IV. RESULTS

The results of the analysis are summarized in Table I.[1] The answer to the main problem appears simple and quite definite. In the vast majority of cases, it will be seen, there is conclusive evidence for a single factor only—the first or 'general' factor. In five cases, however, there are significant indications of a second factor as well. In no case is there acceptable evidence in favour of a third. The average contributions to the total variance are, for the first factor, 0·48, for the second, 0·11, for the third, 0·04. Thus, as with correlations between tests, one 'general factor' almost invariably stands out as responsible for the majority of the variance; more specific factors, though doubtless present, are so uninfluential as to be exceedingly hard to detect.

In discussing certain conclusions, drawn by Mr Cohen ((26), cf. (15), p. 360) and myself, Dr Stephenson has argued that, even when the residuals are too small to be significant, nevertheless, if they fall into a definite bipolar pattern ('system 5' as it has been called ((15), p. 349)), then this pattern of itself demonstrates the presence of a bipolar factor, since such patterns could not be produced by chance. The argument,

---

[1] The correlation tables on which my figures are based, with full details as to saturation coefficients, etc., are given in the second part of my thesis (31). Tables from earlier enquiries have proved difficult to procure; hence researches prior to 1935 are, I fear, very inadequately represented. Owing to the expense of printing or typing the large correlation tables necessary when persons are correlated, most editors, it would seem, have been reluctant to print them in full, and many research students have not included them in their typed theses. I have, however, been able to procure manuscript figures for a few early investigations and a number of the most recent. The references at the end of this paper indicate the enquiries in connexion with which the data were collected. For the calculations I myself am primarily responsible in every case: indeed, comparatively few investigators have attempted factorization, and those who have done so have generally limited themselves to one factor or one small sample.

On the two particular problems that have given rise to the sharpest differences of opinion—olfactory preferences as studied by Beebe-Center and pictorial preferences as studied by Dewar—I have attempted to secure first-hand data of my own. The details are fully described in the first part of my thesis; here I need only say that in most essential points they seem to confirm the views expressed by those writers.

Table I. *Amount of variance attributable to first and second factors respectively in correlations between persons*

		First factor	Second factor
(i)	Types of imagery: unselected children (Burt & Davies(1))	0·32	0·14
(ii)	Types of imagery: selected cases (Burt & Davies(1))	0·28	0·23*
(iii)	Difficulty of Binet tests: examiners (Burt(3))	0·88	0·04
(iv)	Difficulty of Binet tests: children (Burt(3))	0·76	0·08
(v)	Reliability of essay marks (*a*) (Thomson & Bailes(6))	0·63	0·02
(vi)	Reliability of essay marks (*b*) (Thomson & Bailes(6))	0·64	0·06
(vii)	Reliability of essay marks (*c*) (Thomson & Bailes(6))	0·51	0·07
(viii)	Reliability of examination marks (Burt(13))	0·76	0·15
(ix)	Reliability of examination marks (*a*) (Rhodes(13))	0·42	0·08
(x)	Reliability of examination marks (*b*) (Rhodes(13))	0·28	0·13
(xi)	Reliability of examination marks (*c*) (Rhodes(13))	0·49	0·15
(xii)	Reliability of examination marks (*d*) (Rhodes(13))	0·48	0·13
(xiii)	Tactile judgement (Binns & Burt(4))	0·43	0·05
(xiv)	Preferences for pictures: children (Burt(9))	0·73	0·06
(xv)	Preferences for picture postcards: students (Stephenson(15))	0·23	0·16
(xvi)	Preferences for polygonal figures: laymen (Swift. Quoted in (22))	0·48	0·13
(xvii)	Preferences for polygonal figures: students (Swift. Quoted in (22))	0·46	0·07
(xviii)	Preferences for vases: students (Schnittkind. Quoted in (22))	0·33	0·15
(xix)	Preferences for vases: naïve observers (Schnittkind. Quoted in (22))	0·37	0·15
(xx)	Preferences for vases: students of fine art (Schnittkind. Quoted in (22))	0·51	0·08
(xxi)	Preferences for pictures: art critics (Dewar(25))	0·70	0.04
(xxii)	Preferences for landscapes: 1st Series: students (Eysenck & Davies, see (31))	0·31	0·10
(xxiii)	Preferences for landscapes: 2nd Series: students (Eysenck & Davies, see (31))	0·22	0·11
(xxiv)	Preferences for literary extracts (Williams *et al.*(27))	0·55	0·18
(xxv)	Preferences for poetry (Miller. Quoted in (22))	0·34	0·13
(xxvi)	Preferences for nonsense lines (Miller. Quoted in (22))	0·50	0·14
(xxvii)	Preferences for musical extracts (Winter & Wright. Quoted in (27))	0·49	0·11
(xxviii)	Preferences for musical extracts (Wing(33))	0·61	0·12
(xxix)	Preferences for diatonic harmonies (Fischer. Quoted in (20))	0·37	0·15
(xxx)	Preferences for humour (Schonfeld(35))	0·44	0·08
(xxxi)	Preferences for smells (Beebe-Center(10))	0·57	0·08
(xxxii)	Preferences for smells (Stephenson(15))	0·50	0·09
(xxxiii)	Preferences for smells (Davies(31))	0·58	0·07
(xxxiv)	Preferences for school subjects (Burt(8))	0·52	0·16
(xxxv)	Preferences for school subjects (Pritchard(12))	0·47	0·11
(xxxvi)	Preferences for school subjects (Stephenson(14))	0·45	0·18
(xxxvii)	Tests of relation-finding (Cohen. Quoted in (15))	0·24	0·02
(xxxviii)	Intellectual characteristics: normal adults (Burt & Spielman(5))	0·67	0·02
(xxxix)	Intellectual characteristics: normal children (Burt(3))	0·51	0·08
(xl)	Temperamental characteristics: normal adults (Burt & Spielman(5))	0·64	0·01
(xli)	Temperamental characteristics: delinquents and neurotics (Burt(23))	0·60	0·29*
(xlii)	Moods: normal adults (Stephenson(15))	0·46	0·09
(xliii)	Moods: students (Knowles(32))	0·56	0·05
(xliv)	Moods: normal children (Hill(34))	0·49	0·12
(xlv)	Moods: psychotics and normals (Stephenson(17))	0·32	0·24*
(xlvi)	Psychic tempo (Stephenson(16))	0·24	0·16
(xlvii)	Moral reliability (Stephenson(17))	0·41	0·13
(xlviii)	Children's fears (Campbell(28))	0·39	0·22*
	Average	0·484	0·112

* The second factor is statistically significant.

however, appears to overlook the scheme of conditions that must inevitably be fulfilled if any set of intercorrelations, whether arising from chance or not, are to be consistent. These 'consistency conditions' from their very nature tend to produce the appearance of a bipolar hierarchy.[1]

Those, therefore, who have criticized previous investigators for confining themselves mainly to a first or 'general' factor and have maintained the overwhelming importance of the type factors seem so far to have failed in making out their case. To attempt a detailed examination of the several researches in turn would be impossible within the limits of a single paper. But a glance at one concrete problem, which has formed the centre of controversy in a series of three or four articles, will make the general position clearer. Working at Birmingham under Prof. Valentine's supervision, Pritchard (12) endeavoured to study pupils' preferences in regard to school subjects. The 8000 children were classified according to age, sex, and school, and the rankings averaged for the different age-groups somewhat on the lines of previous investigators (8). Although this procedure has hitherto been regularly employed by investigators who have dealt with this and similar problems, Stephenson's conclusions would render it extremely precarious. There must, he argues, be different 'types of preference' within each given age-group; hence it is not justifiable to average all the children of a given age without first applying 'Q-technique' to discover the particular type to which each of the children belongs. Similarly, the attempts by Burt, Bulley, Dewar and others to measure capacity for artistic appreciation are "foredoomed to failure". Capacity for artistic appreciation is not a general factor; hence (as popular opinion rightly maintains) there can be no agreed or common standard. To illustrate these criticisms, he correlates the school preferences of fifteen boys and twenty-two girls, and discovers two distinguishable types, namely, the scientific and the linguistic (14).

Now, Pritchard and his predecessors have never denied the existence of such a bias in many individual pupils. Several have expressly noted it, and have pointed out that it forms the implicit basis of the familiar division into classical and modern sides and into Arts and Science Faculties. Thus, the essential question is, not whether such differences exist, but whether they are large enough, wide enough, and numerous enough to invalidate the general average. An answer can readily be obtained. If we take the tables printed to demonstrate the two contrasted types and partial out the general factor (which Pritchard's average is intended to measure), we then discover that not a single significant residual remains on which the two types can be convincingly based. Nor does the bias extend to more than two or three subjects, or affect those subjects very greatly. We are, indeed, given the order of preference for the children representing the two types; but the correlation between these orders proves to be nearly 0·5. The types, therefore, are not antithetical nor greatly inconsistent: in the 'scientific' type, physics, chemistry, and mathematics go up 2 or 3 places; French and English go down; but for the rest the general order is almost the same as in the 'literary' type. I may add that these conclusions are confirmed by the more recent work of Shakespeare (also from the Education Department at Birmingham) who writes that in his own data "the correlations (between persons) approximate to a chance distribution, and provide no significant divisions like Dr Stephenson's sharply divided groups in pupils of one school" ((19), p. 149).

---

[1] This is demonstrated at length in the appendix to my thesis (31). It is also very evident in Mr Cohen's first research (as cited in (15), p. 360), where only 12 items are correlated for 200 persons. Cf. also (22), p. 349, Table IX, where 88 out of the 105 coefficients fit a bipolar pattern.

Perhaps the acutest conflict, at any rate upon the surface, is between the views of Beebe-Center and Stephenson. Beebe-Center had adopted precisely the same inverted technique as that subsequently employed by Stephenson, and, indeed, had gone further, in that he applied Spearman's tetrad-difference criterion to check the results. Like Pritchard and others, he finds that individual preferences (this time for odours) are correlated highly enough to justify averaging, but discovers no evidence for more than a single factor. Stephenson has repeated Beebe-Center's experiment, and finds at least two factors, the second being relatively specific and bipolar. He claims that Beebe-Center, in describing the results in terms of Spearman's 'r-technique', has overlooked the possibility of a multiplicity of factors, 'which looms large in Q-technique though rare in r-technique'. Unlike Beebe-Center, however, he applies no criterion to determine whether his second ('specific' or 'group') factor is significant or not, which is a little puzzling, since, when correlating tests instead of persons, he always rejects specific factors if they are not corroborated by the tetrad-difference criterion. In Table I, I give the variances for the two factors as obtained from Stephenson's own tables ((15), Tables II and III, pp. 356–8); and here again, after the main general factor has been eliminated, the residuals on which the types are based ('Factor II') prove to be all without statistical significance. It is only fair to add that Dr Stephenson's criticisms are reinforced by more general arguments ((15), pp. 353, 355); but, as I gather from Dr Harsh, these are partly due to an "excusable misunderstanding of Beebe-Center's viewpoint". The number of test stimuli on which the correlations were based is understated; and further it is wrong to suppose that Beebe-Center was seeking to discover "the persons with the most ability", i.e. "with the keenest olfactory susceptibilities"; his general factor was intended to measure (like Burt's and Pritchard's) not individual ability, but individual agreement with the general order. In a reply which he is good enough to forward on Prof. Beebe-Center's behalf, he describes more recent experiments with a Thurstone technique, and adds: "we certainly should not deny the existence of second factors or types...but are withholding judgement until we see if such factors operate in other groups."

This does not mean that Stephenson's very suggestive criticisms can be forthwith dismissed. The opposing views, I think, may easily be reconciled in the following way: we may agree with Stephenson that specific type factors probably exist; indeed, many of those that he mentions have already been indicated by correlations between tests (cf. (20), p. 65). At the same time we may agree with Beebe-Center and Pritchard in believing that, compared with the general factor, their influence is so slight as to be relatively insignificant in most of the experiments so far made. By a special choice of test material or of the individuals tested, we may, no doubt, bring them into prominence. But even so, like all such 'factors', they will be in the first instance mere statistical abstractions; they will be descriptive of the particular correlation matrix, but not necessarily descriptive of the individuals or the traits. To analyse a table is not to analyse persons. We require first to prove that the classification thus suggested is a stable one, and secondly to demonstrate its psychological nature.

The stability of the types may be tested in the following way. Burt has shown that, if two correlation matrices A and B have exactly the same factors, then they must commute; i.e. their products AB and BA must be equal and therefore symmetrical ((20), p. 81). I have given one and the same set of students two different sets of pictures to rank in order of preference, and have correlated the results. The two correlation tables admirably fit the criterion; only one column of products shows deviations from the requirements of symmetry by more than three times the probable error, and that relates to the preferences of one particular person (myself) who deliberately changed the ground of choice in the second experiment.[1]

To determine the nature of the factors, the statistical study must be followed by, and if possible preceded by, a psychological study. Indeed, as a general rule, it is only when the numerical tables are supplemented by introspective data (as in (25)) or by an intensive clinical study (as in (23)) that the psychologist becomes convinced of the existence of the 'type tendencies', and begins to discern their concrete character. As Bartlett has observed: "The practice of separating observers into types may have a definite value, since it may help us to handle a mass of data easily. But we have no right to regard them as psychological starting points, unless we have better reasons for this than the mere classification itself.... Thus, a type having been distinguished, we still have to discuss the interrelation or characteristics by which it has been built up, or, failing that, give adequate reasons for regarding it as innate" ((29), p. 39).

From the data already accumulated, we can, I think, draw many instructive inferences, which may be helpful for further research. Briefly epitomized, the following are the main conclusions that emerge.[2] Most of them can be verified by the reader from a study of the figures given on p. 413, if he is familiar with the general problems and methods of the various investigations (see also Davies (31)).

---

[1] Detailed figures, together with a full description of the method, will be found in my thesis (31); for this part of the investigation I am deeply indebted to the assistance of Mr Eysenck.

[2] Those familiar with Burt's *Notes on Factor-Analysis* will recognize that many of my conclusions follow his, at times almost verbatim. But that seems unavoidable, since (as was to be expected) the more obvious and constant features such as were discernible in his briefer review would necessarily reappear in a wider one. I am specially grateful to him for permission to quote or modify his statements.

V. SUMMARY OF CONCLUSIONS

A. *As regards methods*

(1) Where the persons correlated form a random sample, the method of factor analysis should proceed by an analysis of the correlation matrix *as a whole* ('method *b*'); but where the persons correlated are unavoidably compounded of samples drawn from distinct and discontinuous groups, and particularly where there is independent evidence showing the nature of this grouping, there a separate analysis of the positive submatrices ('method *a*') may prove the more effective procedure. With one or two exceptions, this double rule conforms to the practice of all the investigators in our list.

(2) In all the cases reviewed, the 'factors' which result from the statistical analysis appear to be capable of an intelligible psychological interpretation, *as they stand*. *Prima facie*, therefore, there seems to be no need for "rotation of the axes of reference": nor apparently has any investigator who has analysed correlations between persons attempted such rotation. (This, of course, does not preclude the possibility that rotation may be appropriate to future problems.)

(3) In two important respects the factor analysis of correlations between persons appears *simpler and more direct* than that of correlations between traits:

(*a*) The 'bipolar' factors (factors with negative as well as positive saturation coefficients) have an immediate and intelligible interpretation where correlations between persons are concerned, even if (as many have maintained) this is not so when correlating cognitive tests. Consequently, in correlating persons there is no need to "rotate the axes until there are no negative loadings".

(*b*) There is little reason to believe that any person should have a measurement of precisely zero for any particular factor. Consequently in correlating persons there is equally no need to "rotate the axes until a maximum number of the factor loadings are zero".

B. *As regards results*

(i) *First or general factors*

(1) Except when based on selected material, correlations between persons, like correlations between traits, tend almost universally to display some approximation towards a *hierarchical* arrangement, that is to say, towards an arrangement in which the relative influence of the first or 'general' factor is much greater than that of any other. Hence

the influence of all special (or group) factors becomes exceedingly difficult to demonstrate, and, unless the correlated items are exceptionally numerous, proves nearly always to be statistically insignificant. In forty-four out of the forty-eight researches here examined the first factor will account for all the variance and covariance within the limits imposed by the probable error.

(2) In correlations between persons as in correlations between traits, the saturation coefficients for the first factor are, as a rule, *positive* throughout: only in exceptional cases (usually arising out of homogeneity of persons or test material) are they partly positive and partly negative, i.e. *bipolar.*

(*a*) *Positive or unipolar factors.* (3) When positive, the first factor appears to be descriptive of a tendency *common to the total population* (or at any rate to the total sample) as distinct from more specific tendencies characteristic of particular individuals, types, or subgroups.

(4) In certain investigations this factor may represent nothing but an irrelevant or artificial influence—an intrusive *error or halo-effect,* which needs to be eliminated before the more specific factors can be brought to light. In correlating traits, 'halo-effects' (as is well known) may centre about certain persons; in correlating persons, similar 'halo-effects' evidently gather about certain traits. In both cases the appearance of a 'spurious' general factor is produced. Such a result is particularly noticeable when the examinees are required to assess *themselves.* Traits with attractive names may be ranked high by all; traits with repugnant names may be ranked low by all. Even when the judges are estimating *other persons,* they tend to rank high (or low) for each particular person a trait which ranks high (or low) for the whole human race.

(5) In other investigations this general population factor may be of interest and importance for itself.

(i) As a factor of *reliability,* it indicates that all the judges are tending to estimate or grade one and the same quality and to use one and the same set of criteria.

(ii) As a factor of *preference,* it tends to show that individual likes and dislikes (e.g. aesthetic, ethical, and even scholastic and sensory preferences) are to some extent general or 'objective' (in the sense that they are not solely based on individual peculiarities and so not purely subjective).

In these two cases the superior predominance of the first or general factor justifies the practice of taking an average (or a weighted average) of the individual judgements and regarding this as an 'ideal' measure of the 'true' merits of the examinees or of the test objects. The presence of different mental 'types' does not invalidate this procedure (as has been supposed), because the type factor contributes so small a proportion of the total variance.

(6) In both cases *the first or general factor usually accounts for about half the total variance* (between 30 and 70 %). The remainder of the variance is due, not to any one predominant positive factor or "to an obverse or antithetical type tendency", but to a number of bipolar factors whose influence is diminishingly small.

(*b*) *Bipolar factors.* (7) A bipolar factor appears as the first factor only when the persons or the objects judged form a comparatively *homogeneous* group, e.g. when, in estimating emotional traits, all traits have the same average value, or when, in judgements of preference for pictures, etc., all the pictures are of approximately the same general artistic merit. The selection of a group of items that are homogeneous in regard to the qualities to be judged is in fact equivalent to eliminating by experimental rather than statistical methods the general or positive factor, and so converting what would otherwise be a second factor into the first. Certain recent experiments suggest that this technical device supplies a useful method of procuring significant evidence for the bipolar factor, when the ordinary method (which leaves it to appear as a second factor rather than the first) is likely to fail.

(8) Under these conditions negative as well as positive correlations then appear in the original table. Such a pattern, however, is *exceptional*, and does not (as has been stated) form the general rule. And even so the table still approximates to hierarchical form. It is by no means necessary (as has been erroneously maintained) to introduce a separate specific factor for each positive matrix or for each separate type; if this is done, then the two factors, so far from being independent, prove to be highly correlated.

(9) In virtue of the 'consistency conditions' obtaining among correlations, the appearance of a bipolar hierarchy may arise among correlations whose varying values are *due exclusively to chance.* By itself, therefore, the appearance of such a pattern provides no evidence for the existence of a significant bipolar factor or of antithetical types.

(ii) *Secondary or specific factors*

(1) All factors except the first prove to be bipolar. When genuinely significant such bipolar factors may be taken as indicating the presence of two *antithetical types or tendencies.*

(2) Since in most published investigations an overwhelming proportion of the variance proves to be attributable to the first factor, it is not surprising to find that the saturation coefficients for secondary or specific factors generally have *little or no statistical significance*. The practical corollary seems obvious: when the method of correlating persons is used to discover so-called mental types, then either the number of traits must be very large, or the persons (and in certain cases the tests) must form a homogeneous group. When this is not feasible, then the types and the factors must be elicited, not by correlating persons, but by correlating traits.

(3) In aesthetic preferences, in children's preferences for school subjects, and in temperamental traits, different investigators seem to have reached quite independently much *the same factors or types*. Nevertheless, with rare exceptions, it would seem that their conclusions really rest more upon general observation than upon the statistical evidence which they adduce. The statistical evidence is consistent with their views, but rarely affords any proof.

(4) Where the same group of persons has been tested with different versions of the material and so has supplied two or more matrices of correlations, the symmetry criterion may be used to demonstrate that *the same factors are present in the different matrices*. It appears that the factors reached by correlating persons are by no means necessarily limited by the 'set up' of the test-material on which they are based, but on the contrary are comparatively stable for any given set of persons. Under certain conditions (viz. similar age, attitude, intelligence and variability of the different groups tested) it seems possible to show that they are also stable for one and the same set of tests.

REFERENCES

(1) Burt, C. (1912). *J. exp. Ped.* I, 355–88.

(2) —— (1917). *Distribution and Relations of Educational Abilities*. London: P. S. King and Son.

(3) —— (1921). *Mental and Scholastic Tests*. London: P. S. King and Son.

(4) Binns, H. & Burt, C. (1922). *J. Nat. Inst. Industr. Psychol.* I, 93–8.

(5) Burt, C., Gaw, F. & Spielman, W. (1926). *A Study in Vocational Guidance*. London: H.M. Stationery Office.

(6) Thomson, G. & Bailes, H. (1926). *Forum of Education*, IV, 85–96.

(7) Spearman, C. (1927). *Abilities of Man*. London: Macmillan and Co.

(8) Burt, C. (1931). Board of Education, *Report of the Consultative Committee on the Primary School*. London: H.M. Stationery Office.

(9) —— (1933). *The Psychology of Art*, ap. Burt, C., Jones, E., Miller, E., Moodie, W. *How the Mind Works*. London: Allen and Unwin.

(10) BEEBE-CENTER, J. G. (1933). *Pleasantness and Unpleasantness.* New York: Century Press.

(11) FISHER, R. A. (1934). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

(12) PRITCHARD, R. A. (1935). *Brit. J. educ. Psychol.* v, 157–79, 229–41.

(13) HARTOG, P., RHODES, E. C. & BURT, C. (1936). *Marks of Examiners.* London: Macmillan and Co.

(14) STEPHENSON, W. (1936). *Brit. J. educ. Psychol.* VI, 43–57.

(15) —— (1936). This *Journal,* XXVI, 344–61.

(16) —— (1936). *Char. and Pers.* IV, 294–304.

(17) —— (1936). *J. educ. Psychol.* XXVII, 353–67.

(18) —— (1936). *Psychometrika,* I, 195–209.

(19) SHAKESPEARE, J. J. (1936). *Brit. J. educ. Psychol.* VI, 147–64.

(20) BURT, C. (1937). This *Journal,* XXVIII, 59–96.

(21) —— (1937). *Brit. J. educ. Psychol.* VII, 172–95.

(22) BEEBE-CENTER, J. G. & PRATT, C. C. (1937). *J. gen. Psychol.* XVII, 339–54.

(23) BURT, C. (1938). *Brit. J. med. Psychol.* XVII, 158–88.

(24) —— (1938). *J. Psychol.* VI, 339–75.

(25) DEWAR, H. (1938). *Brit. J. educ. Psychol.* VIII, 29–49.

(26) COHEN, J. I. (1938). *J. ment. Sci.* LXXXIV, 495–512.

(27) WILLIAMS, E., WINTER, L. & WOODS, J. M. (1938). *Brit. J. educ. Psychol.* VIII, 231–49.

(28) CAMPBELL, H. F. (1938). *Brit. J. educ. Psychol.* VIII, 314.

(29) BARTLETT, F. C. (1938). *Remembering.* Cambridge University Press.

(30) VERNON, P. E. (1938). *Psychological Tests of Personal Qualities.* London: H.M. Stationery Office.

(31) DAVIES, M. (1938). "The General Factor in Correlations between Persons." Unpublished Thesis, University of London Library.

(32) KNOWLES, E. (1938). "A Study of Extraversion and Introversion by *Q*-technique." Unpublished Thesis.

(33) WING, H. D. (1938). "Tests of Musical Ability." Unpublished Thesis.

(34) HILL, M. E. (1938). "A Genetic Study of Temperament." Unpublished Thesis.

(35) SCHONFELD, A. (1938). Unpublished work, University College, London.

(*Manuscript received* 15 *June* 1938)

TEMPERAMENT AND THE THRESHOLD FOR FLICKER

By LL. WYNN JONES

I. *The flicker test and mania-melancholia* (p. 422).
II. *The flicker test and introversion-extroversion* (p. 423).
III. *The flicker test and perseveration-non-perseveration* (pp. 423–424).
IV. *The flicker test and ideo-motor tests* (p. 425).
V. *The need of a more comprehensive investigation* (pp. 425–426).
References (p. 426).

The flicker test has been used in the study of melancholia, introversion, and perseveration, but strict definitions of these three terms and their antonyms are not available. Positive statements concerning them or their interrelations must therefore be formulated with care. It will be shown later that lack of such care easily leads to semantic confusion and to gross violations of the rules of Aristotelian logic.

I. THE FLICKER TEST AND MANIA-MELANCHOLIA

As regards melancholia, let us assume that the staff of a large mental hospital were suddenly replaced by a new staff who, without access to the case records, had to label each patient anew. Would the resulting 'melancholia' group coincide with the original, even if typically circular cases were excluded from consideration? Let psychiatrists give the answer. "Psychiatric syndromes are so protean in their manifestations as to be differently classified by competent observers with only slightly different orientations" (1). "We are far from satisfied with our diagnoses" (2). "General formulae are dignified, and diagnostic terms give comfort, but they are verbal symbols which are apt to do violence to the complexity of the facts" (3). Such terms, however, are obviously essential to the psychiatrist. It is the ordinary psychologist who needs to exercise care in their use.

The flicker test was employed by Wiersma (4,5), who found the speed of fusion of colours to be much lower for melancholic than for manic subjects. The present writer (6,7), on the other hand, found no such distinction. It is true that Wiersma included cases of paranoia in his melancholic group, but this is not likely to be the sole cause of the discrepancy between the two findings.

II. THE FLICKER TEST AND INTROVERSION-EXTROVERSION

Washburn *et al.* (8) found that the speed of fusion was actually lower for extroverts than for introverts, but it must be remembered that introversion tests can hardly be regarded as yielding explicit measurements. Guilford (9, 10), for example, found no single factor running through all the items of an introversion test.

III. THE FLICKER TEST AND PERSEVERATION-NON-PERSEVERATION

As regards perseveration, the need of definition is also manifest. A paper by Stephenson (11) specially emphasizes this need. The work of Lankes (12) has served as a beacon light to all subsequent workers in this field. Lankes studied the interrelations of several phenomena and produced evidence that the speed of fusion was related to perseveration.

Quite recently Biesheuvel (13), by means of an ingenious device similar in principle to the stroboscope, also concludes that the speed of fusion yields a measure of perseveration. The present writer has no grounds for doubting the truth of this conclusion. Nevertheless, the aim of this communication is to consider some misconceptions which may obscure the issues when the flicker test is studied in relation to orectic traits.

Biesheuvel devotes more than a tenth of his article to a consideration of Lankes's method and makes the astonishing statement that "Lankes's procedure has been described at length because it was also adopted by Wynn Jones"! It is true that "I adopted several improvements due to Lankes such as descending from a high speed instead of ascending from a low speed", and I still consider that such a procedure yields reliable measurements in the hands of those acquainted with psychophysical methods. Biesheuvel has apparently failed to realize how widely my aims and procedures differed from those of Dr Lankes.

(1) My subjects were patients at a mental hospital. It was therefore necessary to use a simpler technique than that which served Lankes so well with students. An experienced tester like Lankes would not have considered a session lasting forty minutes when testing abnormal subjects.

(2) My subjects were tested individually.

(3) My disks were driven by an electric motor.

(4) I was concerned in comparing the reactions of manic and melancholic subjects in the flicker test, irrespective of whether it was a measure of perseveration. Such being the differences in procedure it is unnecessary for my present purpose to give further consideration to the criticisms of the Lankes technique, although some of them could be refuted. On

reading further I find that "Wynn Jones's contradictory evidence must be rejected, his apparatus and technique not being suitable for the measurement of the threshold for flicker". Now Biesheuvel's efforts to develop "a careful and standard technique" of using the flicker test will be commended by everyone. But he has obviously not realized what my evidence was intended to convey. Unlike Wiersma I found that the flicker test did not distinguish between cases of melancholia and cases of mania. I said nothing to indicate that it was not a test of perseveration. I therefore suggest to Biesheuvel a method of checking my conclusions. Let him try his own method, unless it is too complicated, with melancholic and manic subjects. But the following rules of the game must be observed. He must take a melancholic and a manic group duly selected by hospital doctors and no case must be rejected without reasonable cause. He should preferably not be told to which group a patient belongs until his fusion speed has been determined. If he found a result different from mine he would have some grounds for rejecting my evidence.

Biesheuvel also raises the important question of dark adaptation, but I failed to see any reference to the period which he adopted. Ross (14), in his well-controlled investigation at Yale, kept his subjects for two minutes in a lighted room; then there followed one minute of darkness before the stimulus was lighted. Although I tested each subject with red-green and with blue-yellow disks on five different days I cannot claim to have rigorously controlled the factor of adaptation. I was fully cognizant of its importance, for previous to that time I had probably had more dark-adapted subjects than most psychologists and had on many occasions acted as subject, e.g. as one of the three subjects of Dr Goldschmidt in his *Eigenlicht* researches in Wundt's laboratory. I now turn to the statistical significance of Biesheuvel's results. One weakness he has himself pointed out, namely, that the estimates of perseveration were supplied by only one judge. Moreover, the small sizes of his two groups preclude the drawing of a positive conclusion. His first group numbered only nineteen cases. We are not told how many of these were perseverators and how many non-perseverators. We are told that the difference between the average flicker threshold for the perseverators and non-perseverators was three times the standard deviation. Actually I find it was 2·6. The corresponding figure for his second group of twenty-three cases is given as four. Actually 3·9.

In the absence of any data on the intra-group variability, would any statistically minded investigator attach value to such a test for measuring perseveration in the *individual*? or even for measuring the difference in perseveration between two groups?

IV. THE FLICKER TEST AND IDEO-MOTOR TESTS

In advocating the claims of his test Biesheuvel further writes: "The introverted individual will be much more at ease observing and reporting a subjective experience than performing a task." He is thus contrasting the experience of judging flicker with performance in a motor test of perseveration. He seems unaware of the fact that Pinard, Cattell, and others have produced evidence that it is entirely wrong to connect high and low perseveration with introversion-extroversion (15, 16). For Biesheuvel, melancholia, introversion and perseveration are apparently almost synonymous terms. I agree with him that the flicker test may be a good instance of Spearman's law, but it does not necessarily follow "that the threshold for flicker should provide a better measure of perseveration than the usual ideo-motor tests". This can only be settled by experiment. It is at least significant that since 1914, when I employed motor tests (17) embodying the principle of interference with a well-formed habit (e.g. the reverse stroke test and the "it" test), and the alternation principle (e.g. the ƧS test), dozens of investigators have followed suit and have improved the technique. The reason for the prevalence of such tests is that they produce results and, as they are group tests, they produce quick results. One may, however, hazard the opinion that some of them would prove still more serviceable when applied to one subject at a time as is necessarily the case with the flicker test. There are so many overt movements in the motor tests which cannot be controlled when employed with groups. I am at present attempting to use tests when such movements are kept under control.

V. THE NEED OF A MORE COMPREHENSIVE INVESTIGATION

In conclusion I may say that it is no secret to many that I started a series of individual tests of various kinds about the same time as the group tests mentioned in the *Abilities of Man*. Amongst them were tests of the following phenomena: cold, heat, pain and pressure after-effects, the after-effects of seen movements, after-images and induction effects, e.g. when one tapping rate is followed by a faster. Unfortunately the work was interrupted before a sufficient number of subjects had been tested. It may be of interest to future investigators if I refer very tentatively to a few 'pointers'. There were two kinds of tests which curiously enough did not seem to correlate with the others, viz. tests of the after-effects of seen movements and induction tests.

It was also found that some of the after-effects studied required from the subjects a considerable degree of skill in introspection.

The results of this preliminary testing certainly indicate the need of a more extensive investigation into the interrelations of such phenomena and their bearing on Spearman's law of inertia: *Cognitive processes always both begin and cease more gradually than their* (*apparent*) *causes*. But there still remains the formidable but necessary task which has been recommended by Prof. Spearman, namely, to investigate further the scope of this law with reference to various processes or tasks.

REFERENCES

(1) Masserman, J. H. & Carmichael, H. T. (1938). "Diagnosis and prognosis in psychiatry." *J. Ment. Sci.* LXXXIV, no. 353.

(2) Bond, E. D. & Braceland, F. V. (1937). "Prognosis in mental disease." *Amer. J. Psychiat.* XCIV, 263.

(3) Campbell, C. Macfie (1935). *Destiny and Disease in Mental Disorders, with Special Reference to Schizophrenic Psychoses.* New York: W. W. Norton and Co.

(4) Wiersma, E. (1906). "Die sekundäre Funktion bei Psychosen." *J. Psychol. Neurol., Lpz.*, VIII, 1.

(5) —— (1932). *Lectures on Psychiatry.* London: H. K. Lewis.

(6) Jones, Ll. Wynn (1928). "An investigation into the significance of perseveration." *J. ment. Sci.* LXXIV, 653.

(7) —— (1929). "Individual differences in mental inertia." *J. Nat. Inst. Industr. Psychol.* IV, 282.

(8) Washburn, M. F., Hughes, E., Stewart, C. & Sligh, G. (1930). "Reaction time, flicker and affective sensitiveness as tests of extraversion and introversion." *Amer. J. Psychol.* XLII, 412.

(9) Guilford, J. P. (1933). "An examination of a typical test of introversion-extraversion by means of the method of similar reactions." *J. Soc. Psychol.* IV.

(10) Guilford, J. P. & Guilford, R. B. (1934). "An analysis of the factors in a typical test of introversion-extraversion." *J. Abn. Soc. Psychol.* XXVIII, 377.

(11) Stephenson, W. (1934). "An introduction to so-called motor perseveration tests." *Brit. J. educ. Psychol.* IV, 186.

(12) Lankes, W. (1915). "Perseveration." *Brit. J. Psychol.* VII, 387.

(13) Biesheuvel, S. (1938). "The measurement of the threshold for flicker and its value as a perseveration test." *Brit. J. Psychol.* XXIX, 27.

(14) Ross, R. T. (1936). "The fusion frequency in different areas of the visual field." *J. Gen. Psychol.* XV, 133, 161.

(15) Pinard, J. W. (1932). "Test of perseveration." *Brit. J. Psychol.* XXIII, 114.

(16) Cattell, R. B. (1933). "Temperament tests." *Brit. J. Psychol.* XXIV, 20.

(17) Jones, Ll. Wynn (1915). "Perseveration." *Rep. Brit. Ass.* p. 698.

(*Manuscript received* 8 *December* 1938)

A FURTHER NOTE ON THE DEGREE OF VARIATION IN THE SCORE OF A MOTOR TEST OF PERSEVERATION

BY JANE DARROCH

IN a discussion with the writer of her previous paper, "An Investigation into the Degree of Variation in the Score of a Motor Perseveration Test" (this *Journal* (1938), XXVIII, 248), Mr Babington Smith pointed out that the writer in her previous article, in order to estimate whether the score for the w test was sufficiently constant for the test to be of value, used as a criterion that the scores of high and low perseverators should not overlap with those of medium ones, and that this criterion was a very exacting one, and at the same time likely to be inexact, since the extreme cases of distributions tend to be unreliable, and that it was inefficient, since only a small part of the information collected was used. Mr Babington Smith proved, by the use of Fisher's analysis of variance, that the differences between the average scores for all the trials included in the tables in the previous article, that is, for about thirty trials, were greater than was due to chance; and pointed out that, since that was so, some importance could safely be attached to the differences between scores. He raised the question whether the same criterion was satisfied by the results of the first five, or any five consecutive, tests taken by the subject, this being a more manageable number of tests to give.

The writer accordingly applied the criterion of Fisher's analysis of variance to the average score for the first five trials, scored by the methods $(X_1+X_2)/Y$ and X_1/X_2, and to that for the trials from the sixteenth to the twentieth inclusive, scored by the same two methods. The score $X_1/X_2+2X_1/Y$ was ignored, because it is so like X_1/X_2 both in variability and in the rank order in which it places the subjects, that anything true of X_1/X_2 may be assumed to be true of it too. Only those eight subjects were included in the calculation who had had no previous practice at the test. For this number of subjects and of trials, there is a 1 % probability that Z exceeds 0·5481 as a result of pure chance. It was found that the mean scores for the first five trials did differ more than was due to chance when they were scored by the method X_1/X_2, Z being 0·7503; and that the same was true of the trials from the sixteenth to the twentieth inclusive, whichever method of scoring was used, since Z was 0·6506 for

$(X_1+X_2)/Y$, and 1·0614 for X_1/X_2. The average score for the first five trials did not satisfy the criterion when the method $(X_1+X_2)/Y$ was used, Z being in this case 0·4362.

It thus appears that, at least when the method of scoring X_1/X_2 is used, the differences between mean scores for five successive trials is sufficiently great for some reliance to be placed on the results of the test. As every trial was included in the calculation, whether any disturbing influence was present or not, except one trial which was mistimed, it also appears that disturbance by such factors as noise and fatigue can be safely ignored unless it is very gross.

As already noted, the rank order of the subjects changes considerably between the earliest trials and the later ones. For the score X_1/X_2, it was F, A, N, B, E, H, M, I for the mean score of the first five trials, and A, F, M, H, E, N, B, I for the mean score of the trials from the sixteenth to the twentieth. Thus, although each of these scores differentiates the subjects satisfactorily, they do not measure precisely the same qualities. The question remains: What is the cause of the difference, and which score is the more likely to yield a measure of some temperamental quality? It is noteworthy that all the statistically significant correlations so far obtained between perseveration tests, scored by the ratio method, have been between first and second trials.

PUBLICATIONS RECENTLY RECEIVED

Experimental Psychology. By ROBERT S. WOODWORTH. New York: Henry Holt and Co. 1938. Pp. xi+889. $3.80.

There is at the present day a widespread belief that experimental psychology has failed to fulfil its early promise as a scientific discipline. An alarming sense of disillusion is abroad, penetrating even into our laboratories. Our students acquire an ever-growing distrust in the confident enthusiasm of the early pioneers. Our 'qualified psychologists' only too often qualify their psychology with a tolerant resignation to its manifest shortcomings. Has, then, the mountain of Leipzig in labour brought forth but a mouse—or at best, a humble white rat?

Prof. Woodworth's recent book, despite its many and obvious qualities, will do nothing to combat this prevalent attitude of disillusion. Every psychologist feels intense gratitude to the founders of his science—to Fechner, to Wundt, to Ebbinghaus, for their untiring efforts to establish an experimental science of mind. But many of us find our admiration of these men not untinged by envy. Cannot we too develop a like spirit of optimism and confidence—if not in ourselves, then at least in our students? Prof. Woodworth's book does not help us here. It is no monument to the passionate belief held out by the pioneers of experimental psychology in the future of their science. Far rather is it a mausoleum of their methods. The author is dominated by the *means* devised by the pioneer experimentalist. What we require to-day is a reaffirmation of his aims and ideals.

Experimental Psychology opens with a chapter on Memory. This contains little else than classification of the 'exact methods' devised by Ebbinghaus and elaborated at great length by his followers. The next two chapters, on Retention and Memory for Form respectively, are replete with facts, both quantitative and qualitative. But the underlying theory is of the sketchiest. Woodworth's assumptions are the familiar postulates of every trace theory, and they are couched in the vaguest terminology of function. Indeed Woodworth's concept of the memory trace is about as 'precise and scientific' as is Ebbinghaus's technique of the nonsense syllable. It would be unfair to allow dissatisfaction with theory to blind one to the obvious merits of these three chapters on memory. There is an immense amount of material here, and the treatment is expressly adapted to the needs of the student who must perforce familiarize himself with the very great mass of controlled experimental data. The presentation, moreover, is both clear and intelligent throughout.

Chapter v deals with the Conditioned Response. This is far more constructive than any of the preceding chapters. The author has made a real effort to grasp, and clearly to set forth, the implications of conditioning for psychology as a whole. The latter half of the chapter comprises, indeed, one of the best general discussions which we have as yet had of the relation of the conditioned response to other forms of behaviour, and to the theory of learning in general. It is marred only by a rather crude and anthropomorphic 'sign theory' of the conditioned stimulus. Chapter VI, on Maze Learning, is also interesting. Prof. Woodworth takes, very properly, as his problem *how the maze is learnt.* His emphasis throughout is upon accomplishment rather than explanation, and the empirical treatment adopted is highly successful. Moreover, Prof. Woodworth's presentation of the experimental studies in historical sequence is of enormous assistance in enabling the reader to gain some sort of an orientation in a field where so much of the contemporary discussion is on the level of highly specialized and stylized theory. And the value of the chapter for the non-specialist reader is further enhanced by the excellent comparative treatment of maze-learning in animals and men. Chapter VII, on Practice and Skill, is competent though not very stimulating. A good deal of the standard objective data is summarized, but there is practically no discussion of the really vital theoretical issues underlying it.

There is likewise no mention of the intensely significant work of Michotte and his pupils on the 'morphology of movements'. This is a notable omission. Chapter VIII and IX deal with the Transfer of Training and Economy and Interference in Learning. There is an enormous amount of experimental material in these chapters, but one cannot avoid the suspicion that the problems are not really worthy of so detailed a treatment. The same suspicion attaches in even more marked degree to the immediately succeeding chapters on some aspects of feeling and emotion.

Chapter XIV, on Reaction Time, is of interest primarily from the standpoint of psychological history. Problems of reaction time are scarcely of central importance to-day, but one may venture to suggest that at no time will the psychologist be in a position wholly to dispense with these measurements. Prof. Woodworth's chapter deserves to become the standard reference for all who have to make use of reaction times in an experimental or clinical capacity.

Chapter XV is called (rather misleadingly) Association. We may hasten to point out that it contains no discussion whatsoever of the celebrated principles. The chapter is entirely devoted to systematizing the results of innumerable experimental studies of word-association. Its value as a summary is considerable, and this chapter, also, is doubtless destined for the fate of a standard reference.

Chapters XVI–XVIII are concerned with Psycho-physics. Here the author's obsession with the methods of Fechner and Wundt emerges most clearly. These methods may, it is true, have the 'disciplinary value' grudgingly accorded them by James, but what is required far more urgently at the present day is a redefinition of the *problems* of psycho-physics. Prof. Woodworth's discussion of psychological measurement, for all its superficial plausibility, manages to evade every serious criticism of the traditional arguments. And he is quite certainly insufficiently critical of Fechner on the psychological side. These chapters, admirable as in many respects they are, clearly point to the past rather than to the future. The present-day student may derive considerable instruction in method from a careful study of this section of the book. But for his inspiration he must needs seek elsewhere.

Chapters XIX–XXII deal with the Senses. The first of these, on the skin senses, is careful and detailed enough, though the recent and highly significant researches of Woollard are entirely omitted. Chapter XX, on Smell and Taste, is neither as complete nor as up to date as one might wish, and no recent work on threshold determination is quoted for either sense. Chapter XXI, on Hearing, is a brave effort to do justice to a field which is daily becoming more complex and specialized. The expert will note some inaccuracies of fact, and he may well be critical of Woodworth's theoretical interpretations. European workers, moreover, may resent the somewhat scanty treatment accorded to their contributions. It is further to be regretted that Prof. Woodworth has not seen fit to include any discussion of the implications of *Gestalt* psychology for auditory theory. Chapter XXII, on Vision, gives a fairly accurate general discussion of the psychological phenomena. It is, however, somewhat disjointed in tone, and the reader must remain content with a rather superficial evaluation of contemporary theory. The author indeed appears to favour the Ladd-Franklin theory of colour vision, despite the weight of evidence against it. Eye Movements are treated with great thoroughness in Chapter XXIII, though the space devoted to this topic may seem to some excessive. Some problems of visual perception are dealt with in Chapters XXIV–XXVI. In the first of these, the author discusses the appreciation of colour and brightness. There is a very good summary of the principal experiments on brightness constancy, and the value of the chapter as a whole is enormously enhanced by the numerous and excellent diagrams of laboratory set-ups. Theory as usual takes second place, though it might seem to some of us that Koffka's ingenious speculations deserve a more thorough and detailed examination than they actually receive. In the following chapter, the organization of two-dimensional fields receives some consideration. This is a highly eclectic chapter, but it has value as an impartial summary of researches on figure and ground, contour and shape, and the conditions of 'organization' in general. Six theories of perceptual organization are listed at the outset, but little attempt is made in the remainder of the chapter to evaluate these theories in terms of the available data. Nevertheless, the chapter as a whole should be of great value to the student.

and the exposition is clear and straightforward throughout. Chapter XXVI, on Visual Space, is notable for an exceedingly able presentation of the principal conditions making for tri-dimensional experience. In a field such as this, probably the most that the writer of a text-book can do is to set out the data clearly and to assess as far as is possible the relative parts played by the various visual and non-visual 'cues' in the perception of distance. And in such a task as this Prof. Woodworth excels.

Chapter XXVII is concerned with Attention. There is much to be said for the resurrection of this concept in a modern text-book. At last psychologists are beginning to realize the absurdity of a psychology in which the term 'attention' is exclusively used in the instructions given to their subjects. But there is disappointingly little that is new in Woodworth's treatment. Our old friends—span of apprehension, shifts and fluctuations, 'doing two things at once'—feature once again. But that is all. If the notion of 'attention' is to reinstate itself honourably in psychological discussion, it must hold forth promise of greater fruitfulness than this.

Chapters XXVIII–XXX deal with Reading, Problem Solving Behaviour, and Thinking respectively. The last two chapters are quite adequately done, though a specific lack would seem to us the absence of any account of the relations of language and thought, and of language itself. But these two chapters do bring together a mass of experimental data bearing upon the thought processes, and taken together, they furnish an excellent introduction to the experimental study of the higher mental processes in general.

Experimental Psychology, despite its author's long-standing and distinguished connexion with American 'Functionalism', is not written primarily from the standpoint of biological science. The heritage of Fechner and Titchener, of the physicist and the formalist, lies heavy upon almost every page of this long and significant book. The really vital problems of present-day psychology—the problem of levels of functional integration, the problem of cortical function in learning, the problem of 'sensory transfer' and the generalization of response, the problem of 'schemata' and the co-ordination of reaction, and, most pressing of all, the problem of motivation in the widest sense—these find no place at all in Woodworth's text. Even the one outstanding theoretical movement of the present century, the *Gestalt* psychology, is treated throughout more as a collection of unrelated suggestions than as a well-organized body of fact and theory. We may in all sincerity congratulate Prof. Woodworth on his diligence, his competence and his erudition. But we find it hard to sympathize with his choice of the material in a book which must surely become for many years the senior text-book in the English language.

Prof. Woodworth is a brave man. And he deserves the profound gratitude both of teachers and of students in psychology. His significant book will long be remembered, if only as the epitome of a phase in which psychology passed through a period of arrested development—a kind of 'plateau' in its progress. And when all is said and done, Prof. Woodworth has achieved more than any other single man since Hermann Ebbinghaus. *Experimental Psychology* is destined for a place on the shelf beside the *Elemente*, the *Grundzuge*, and the *Principles*. O. L. Z.

General Psychology. By W. J. H. SPROTT. London: Longmans, Green and Co., Ltd. 1937. Pp. vii+446. 7*s*. 6*d*. net.

Psychology for Everyone. By W. J. H. SPROTT. London: Longmans, Green and Co., Ltd. 1937. Pp. xi+446. 8*s*. 6*d*. net.

These two books are identical except for their prefaces. The first is addressed to the student *and* the general public, and the second to the general public alone; the student therefore has the privilege of buying almost the same material for a shilling less. It is to be hoped that this privilege will be very widely exercised, for as a starting text-book in the subject this is a very good and comprehensive volume. It demands, indeed, the exercise of considerable intelligence on the part of the reader. Mr Sprott does not shrink from controversial issues, he does not disguise them, and he presents many sides to most problems, often in an exceedingly concise manner. Since, however, nobody is likely to make any headway in psychology unless he is forced to think for himself, all this is to the good.

The book has an original and rather provocative beginning. Everybody knows that psychologists are constantly taunted with persistent internal disagreements. Mr Sprott wisely faces this straight away. He asserts that the differences between 'schools' of psychologists are often much less important than the differences of label would suggest. He states what the main schools are, proclaims his adherence to none of them, and claims that all good psychologists will change their point of view from that of one school to that of another if their facts and problems seem to indicate that this is desirable. It follows at once that data and problems must be collected from all possible sources, and to this principle the book remains true to the end.

All the early chapters deal with varied forms of conational activity. Since these constitute a field in which experimental research is necessarily somewhat backward, the student is, from the beginning, plunged into a consideration of theoretic interpretation. This must not be taken to mean that Mr Sprott has little regard for facts, but only that the first sorts of facts that he faces are rather general observational ones than of the experimental kind. He writes of The Springs of Action, The Social Environment, The Innate Constitution of Man, and goes on to consider Development, Conflict and Psychopathology. The regard for theoretical construction continues in his chapters on Other People and the Self, and it is only when, in Chapter XII, he arrives at a study of The Perceptual Field, and then goes on to a brief study of sensational reactions and to a longer consideration of the higher mental processes, that he shows himself thoroughly well-equipped to lead the student through a mass of experimental material.

In fact there is much to be said for this order of presentation of psychological material. Students who have not already had a reasonably good training in scientific method in other fields are much more likely to be interested and efficient if they are induced first to think psychologically about the events and problems of daily life and current discussion. It is very difficult to imagine that any student can read this book through, or that any member of the general public can dip into its pages, without being held and interested and helped to think in a genuinely psychological manner. If it has any outstanding point of difficulty it is that rather too many pieces of scattered research are mentioned with inadequate comment. Considering the condensation that is achieved it is remarkably accurate. It is throughout alive and attractive and Mr Sprott is warmly to be congratulated upon his achievement.

Explorations in Personality. By the Workers at the Harvard Psychological Clinic, H. A. MURRAY *et al.* Oxford University Press: Humphrey Milford. 1938. Pp. xiv+761. 42*s*. net.

It would be difficult to overrate the value of this study of personality which has been carried out at the Harvard Psychological Clinic. Forty-eight men of undergraduate age were studied (twenty-four intensively) by about thirty psychologists using a large number of different methods of approach—clinical, observational and experimental. The results for each subject were then collated by a 'Diagnostic council' of five, and from these the nature of his personality was deduced and described in terms of a previously established scheme.

This scheme was based upon two main procedures: (*a*) A list was drawn up of the main 'needs' or drives—achievement, dominance, aggression, etc.—together with certain other important personality variables—emotionality, various tendencies, comprising introversion and extraversion, the nature of the super-ego and the ego-ideals etc.—The subject was graded on each of these variables from observations, answers to questionnaires, etc. (*b*) A 'unity thema' and various subordinate themas were worked out for each subject, mainly from his autobiography, from clinical observation and from the phantasies which appeared in various 'imaginal productivity' tests ('thematic apperception', ink-blot, story writing, 'musical reverie', etc.). This 'unity thema' was "the key of the individual's unique nature"..."a compound of interrelated...dominant needs" and their relation to the important stimulus situations of the subject's childhood and youth.

Though criticism may be levelled at the particular drives and themas selected, there is no doubt that they afford a very valuable and suggestive means for the combined analysis and synthesis of personality. The terminology used is sometimes to be deplored. It may also be objected that the psychoanalytic interests of four of the five members of the Diagnostic council have given an unduly Freudian bent to the themas; and indeed it would seem that the 'Case of Earnst', quoted in full, could be equally well interpreted without recourse to Freudian hypotheses. Prof. Murray, however, has in the Conclusions forestalled and to some extent met criticism on these grounds.

But it is particularly with regard to method that the authors and experimenters, especially Prof. Murray, are to be congratulated; firstly, for the collaboration of experimentalists, psychometrists and clinicians, and their attempts to dovetail their results to produce a coherent whole, instead of merely contradicting one another (this collaboration extends to the writing of the book); secondly, for the study of each individual as far as possible as a whole, rather than the investigation of isolated functions in a variety of individuals. It is to be hoped that this study will be the forerunner of others; though it may be difficult to find a similar group of psychologists who can collaborate without losing their individual outlooks and methods of approach.

The high price of the volume is to some extent necessitated by its great length and fine production; but it may unfortunately prevent many of those who would most profit from its study from buying it.

The Psychology of Social Movements: a Psycho-analytic View of Society. By PRYNS HOPKINS. London: George Allen and Unwin, Ltd. 1938. Pp. 284. 10*s.* 6*d.* net.

Dr Hopkins summarizes the six principal human needs as: (1) Enjoyment of knowledge, (2) Enjoyment of sensory naturalism, (3) Enjoyment of inner peace, (4) Enjoyment of the family, (5) Enjoyment of material means, and (6) Enjoyment of safety from violence. He then considers the means of fulfilment of these needs and the nature of the social and psychological forces which cause their frustration. The point of view from which these problems are considered is throughout psychoanalytical. The author considers that the habit of smoking tobacco plays a large part in the frustration of man's efforts for happiness.

The Mental Health Services of Oxford City, Oxfordshire and Berkshire. By E. F. PINSENT. Oxford: the University Press. 1937. Pp. 87. 2*s.* net.

This quite admirable report on the Oxford Mental Health services is an advance section of the Oxford Survey of Social Services (now published) carried out under the auspices of a Standing Committee of Barnett House. The report describes clearly the Statutory obligations of the Mental Health services, and an investigation of the steps taken to carry them into practice. There have been many valuable beginnings in Oxford City, in spite of lack of encouragement—not to say actual impediment—from the Central Authority. Oxfordshire and Berkshire are more backward. The following statements are of particular interest to psychologists: "We are of opinion that it needs a psychologist visiting these (retarded) classes periodically and examining the children with the class teacher in order to arrive at a valid conclusion as to the most appropriate treatment. A psychologist would also be able to give the teachers information and advice which we think they would greatly appreciate." "It is... unfortunate that all regular observation ceases when a boy or girl leaves school, for it is just at this period that (psychological) help is often necessary to meet the troubles of adolescence." "Steps should be taken to establish a thorough and systematic periodical search in the elementary schools for mentally defective children, and then to provide a special school with suitable and sufficient accommodation for them. We think this search should be carried out by a medical psychiatrist who has special knowledge of teaching, or by an educational psychologist." "There can scarcely be two opinions as to the necessity of having one man in each (Local Government) district with special qualifications, whose opinions can be taken as authoritative, both as regards diagnosis and appropriate treatment, and whose duty it would be to co-ordinate the mental services. He should be a trained psychiatrist, and not only one of the medical staff who has picked up a desultory knowledge of the subject."

Health and Unemployment. By LEONARD C. MARSH, A. GRANT FLEMING and C. F. BLACKLER. Oxford University Press: Humphrey Milford. 1938. Pp. xxv+243. 10*s*. 6*d*. net.

This is a detailed and interesting study of the effect of unemployment upon the physical health of juveniles and adults, and the families of the latter, in Montreal. Unfortunately, it was not found possible to follow up the few suggestive hints as to the psychological outlook of the unemployed, and the effect of unemployment upon their mental health. Useful suggestions are made as to improvements in nutrition and health services.

A Visual Motor Gestalt Test and its Clinical Use. By L. BENDER. Research Monographs No. 3. New York: American Orthopsychiatric Association. 1938. Pp. ix+176.

This interesting book gives an account of the clinical use of a test of copying visually perceived figures. The figures used, eight in number, are of the line and dot variety used by Wertheimer to demonstrate *Gestalt* factors in perceptual organization, and are chosen from the examples given by him in his paper "Untersuchungen zur Lehre von der Gestalt" (1923). These are presented to the patient with the instruction to copy them. In attempting to do so, it is found, he will display forms of breakdown of visual, motor, and other functions which are characteristic of his disorder. In some cases the breakdown involves specifically *Gestalt* factors of organization; in others psychopathic tendencies, such as substitution, fractionation, and elaboration, are manifested. The forms of failure to carry out the test correctly which are found in various organic and psychogenic disorders are described, and the value of the work is much increased by the inclusion of material relating to normal children and adults, with norms for various ages. Dr Bender has opened up a line of clinical research which may come to possess considerable importance, even though the assumptions underlying it have to be modified in the course of time.

Constructing Tests and Grading. By H. D. RINSLAND. London, etc.: George G. Harrap and Co., Ltd. 1938. Pp. xvi+323. 10*s*. 6*d*. net.

This book contains a thorough and very detailed account of 'short answer' tests and examinations of every kind and in every subject, and of the points to be observed in constructing new tests. To those who use this kind of test it will be of value. The detailed discussion of methods of scoring the usual type of examination question and of grading pupils from its results should be of interest to those in charge of large public examinations. But a teacher examining a single class should approach it with caution, since he cannot assume that the performances of a small sample of children will show a normal distribution. From the author's further dictum that "it is only through the experience and improvement of our measuring instruments of to-day, in the simpler phases of human learning, that we can push forward our scientific frontiers to a more accurate and comprehensive description of the personality of the whole child—the new centre of the educational world", we must, however, beg leave to dissent.

Zwillinge und Schule. Von HERBERT GRAEWE. Erfurt: Kurt Stenger. 1938. Pp. 132. RM. 3.60.

This monograph is a study of identical and fraternal twins. It shows how the former are not only physically alike, but also alike in personality, in interests, and in performance in various school subjects. The fraternal twins, however, show marked differences in these respects. The resemblances and differences between the children's drawings are particularly interesting.

Progressive Relaxation. By EDMUND JACOBSON. Illinois: University of Chicago Press; England: Cambridge University Press. 1938. Pp. xvii+494. 22*s.* 6*d.* net.

The first edition of this study was published in 1929 and it attracted considerable attention. The present edition describes new methods for measuring muscular contractions and general nervousness, and purports to describe also relatively exact ways in which the amounts of mental activities present in a patient at a given time may be estimated. It devotes special attention to electro-physiological measurements and technique, and maintains that by the use of these it may be conclusively demonstrated that "during progressive muscular relaxation imagery and thinking processes dwindle and disappear". The clinical significance and therapeutic use of the author's methods, both in a general and in a specific sense, are fully discussed, with some case material. A full bibliography is appended.

Although the book may, at various points, run rather beyond the evidence which is presented, it is a most serious study and deserves careful critical attention and test.

Studies in the Philosophy of Religion. By A. A. BOWMAN. Edited with a memorial introduction by NORMAN KEMP SMITH. London: Macmillan and Co. 1938. Two vols. Pp. xlviii+423; xiii+438. 30*s.* net.

Although these two volumes are, as the title indicates, addressed mainly to philosophers, they contain much that will be of interest to any thoughtful psychologist who has a bent towards critical theory. The second part, in particular, which deals rather fully with *The Anthropological Picture of Religion*, will be found to possess psychological significance, and also the discussions dealing with so-called naturalistic ethics. On the other hand the late Prof. Bowman does not display any detailed knowledge or interest in recent psychological contributions to the study of religious phenomena.

Über Sinn und Sinnlosigkeit des Lebens. Von MARGARETHE V. ANDICS. Wien: Gerold und Co. 1938. Pp. iii+175.

In Vienna, those who are rescued from attempted suicide are (or were) taken to the University Clinic for Psychiatry and Neurology, for observation and treatment. A hundred of these cases were questioned by the author about their life histories and their reasons for suicide, with a view to discovering what it is that makes men feel that their lives are no longer worth living; how their existence becomes meaningless to them, and they seem to be of no use to anyone or for anything. Though the method adopted is psychological, the author's bent and her conclusions are mainly philosophical in nature.

Dasein und Wirklichkeit: Eine Einführung in die Seinslehre. Von OTTO JANSSEN. München: Ernst Reinhardt. 1938. Pp. 351. RM. 8, Leinen RM. 10.

Seele als Äusserung: Untersuchungen zum Leib-Seele-Problem. Von PAUL HELWIG. Leipzig und Berlin: G. B. Teubner. 1936. Pp. 124. RM. 4.20; f. Ausland RM. 3.15.

Neue Dialoge zwischen Hylas und Philonous. Von H. A. WIMMER. Heidelberg: Carl Winter's Universitätsbuchhandlung. 1938. Pp. viii+154. Kart. RM. 7.

Drei Vorträge über Philosophie und Parapsychologie. Von JOHANNES J. POORTMAN. Leiden: A. W. Sitjhoff's Uitgeversmij N.V. 1939. Pp. 76.

FIFTH INTERNATIONAL CONGRESS FOR THE UNITY OF SCIENCE

PRELIMINARY ANNOUNCEMENT

THE Fifth International Congress for the Unity of Science is to be held at Harvard University, September 5 to 10, 1939.

The theme of the Congress is *Logic of Science*; interest will centre upon the relation of the concepts, laws and methods of the various sciences. Attention will be devoted to general problems connected with the unification of science, and, in particular, with the logic of the physical sciences, the relation of the physical and biological sciences, and the relation of the biological and socio-humanistic sciences. There will also be a number of special sessions and symposia concerned with special problems and fields.

The International Organizing Committee is composed of R. Carnap, P. Frank, J. Joergensen, C. W. Morris, O. Neurath, H. Reichenbach, L. Rougier, L. S. Stebbing. The American Organizing Committee is composed of E. T. Bell, P. W. Bridgman, R. Carnap, A. Church, M. Cohen, J. Dewey, H. Feigl, L. J. Henderson, C. J. Herrick, C. L. Hull, E. V. Huntington, H. S. Jennings, W. Köhler, S. K. Langer, C. H. Langford, K. S. Lashley, V. F. Lenzen, C. I. Lewis, R. S. Lillie, A. O. Lovejoy, R. M. MacIver, W. M. Malisoff, W. P. Montague, C. W. Morris, E. Nagel, F. S. C. Northrop, W. V. Quine, H. Reichenbach, E. Sapir, G. Sarton, H. Schultz, E. C. Tolman, W. Weaver, L. Wirth.

Professor P. W. Bridgman is Chairman, and Dr W. V. Quine is Secretary of the Committee on Arrangements at Harvard. The Congress is sponsored by the International Committee of the Congresses for the Unity of Science, by the International Institute for the Unity of Science, and, in America, by the American Association for the Advancement of Science, the Philosophy of Science Association, the Association for Symbolic Logic, and the American Philosophical Association.

A series of twenty monographs, entitled *Foundations of the Unity of Science* (and constituting the first two volumes of the *International Encyclopedia of Unified Science*), is now being issued by the University of Chicago Press, and helps to provide a background for the Congress. Three monographs have already appeared, and it is hoped that all twenty will be in print by the opening of the Congress.

PROCEEDINGS OF THE BRITISH PSYCHOLOGICAL SOCIETY

1938

GENERAL MEETINGS

January 21, 1938. General Meeting, followed by Lecture. (See under LECTURES.)

March 11, 1938. General Meeting, followed by Lecture. (See under LECTURES.)

March 31–April 4, 1938. EXTENDED GENERAL MEETING AT ST ANDREWS.

March 31. Introductory Business Session.

April 1. 1st Session:

"Social Factors in Laboratory Experiments." By B. BABINGTON SMITH.

"Explanations and Demonstrations of a Vocational Test Battery." By H. PALLISTER and W. O'D. PIERCE.

"An Analysis of Vocational Guidance by a Juvenile Employment Bureau." By E. L. TRIST.

"Teaching Method and Subject Preferences: A Quantitative Study in Individual and Social Psychology." By BRENDA ASHBY.

"The Application of Psychoanalytic Methods to the Analysis of a Culture." By P. L. GOITEIN.

"General Review of the above papers and of the work of a Social Research Team." By O. A. OESER.

2nd Session:

"A Psychological Basis for Squint Treatment." By G. H. GILES.

3rd Session:

"The Effects of Errors of Measurement on Correlation Coefficients." By R. H. THOULESS.

"The Effects of Veiling Glare on Apparent Size and Distance." By R. W. PICKFORD.

April 2. 4th Session:

Symposium on "Aggression".

Papers by: W. R. D. FAIRBAIRN, ADRIAN STEPHEN, KARIN STEPHEN, A. W. WOLTERS.

5th Session:

"The Value of Subjective Drawings for the purpose of Psychological Demonstrations." By H. G. BAYNES.

April 2. 6th Session:

"Methodological Implications of Recent Trends in Psychometry." By Wm. Stephenson.

Opening of Discussion on the Terman Revision of the Stanford-Binet Test. Papers by: P. E. Vernon, D. Kennedy-Fraser.

7th Session:

Discussion on the Terman Revision of the Stanford-Binet Test (continued). Paper by: C. A. Simmins.

April 3. 8th Session:

"The Incidence in School Children of 'Word-Blindness'." By M. Macmeeken.

"A Scheme for Vocational Guidance in an Education Area." By A. Macdonald.

"Studies in the Mechanism of Speech." By R. Curry.

9th Session:

"Some Fundamental Points in Stout's Theory of Perception." By J. N. Wright.

"Psychology in the Service of the Next War." By Pryns Hopkins.

10th Session:

Open Session for Discussion.

June 25, 1938. General Meeting at Christ Church, Oxford, and visit to the University of Oxford Institute of Experimental Psychology.

Papers:

"Experiments on the Perceptual Processes among a Primitive People." By M. Fortes.

"General Trends and Individual Differences in the Knee-Jerk." By Wm. Stephenson.

"Some Psychological Principles in Christo-Therapy." By Marcus Gregory.

"The Psychological Bases of Ethics, National and International." By Wm. Brown.

October 22, 1938. General Meeting at Bethlem Royal Hospital.

Papers:

"The Terman-Merrill Revision and Adult Mental Patients." By J. D. Copp.

"An Experimental Investigation of Some Theories of Temperament Types." By E. G. Mountford.

"Psychomotility." By R. J. Bartlett.

"Mental Type and Body Measurement." By Bernard Armitage.

November 18, 1938. Discussion Meeting.

Discussion on "Psychologists and Culture" opened by T. H. Pear.

November 26, 1938.	Joint Meeting with the Institute for the Study of Animal Behaviour. Papers: "Relative Strength of Two Drives in the Black-headed Gull." By F. B. Kirkman. "Grouping and Courtship in Immature Black-headed Gulls." By Miriam Rothschild. "Further Experiments on Drives in Rats and their Physiological Basis." By B. P. Wiesner.
December 10, 1938.	Annual General Meeting. Papers: "Studies of Changes of Interest." By C. A. Mace. "Psycho-Rheology in Industry." By G. W. Scott Blair. Demonstration: "Apparatus for investigating Extra-Sensory Perception." By G. N. Tyrrell.

LECTURES

January 21, 1938.	"Planning for Vocational Guidance." By Alec Rodger.
March 11, 1938.	"The Diagnostic and Therapeutic Value of Play in Childhood." By Grace Calver.

SECTIONAL MEETINGS

Medical Section

January 26, 1938.	"Relationship of Consciousness and the Functional Activities of the Organism." By J. E. Marcault.
February 23, 1938.	Address from the Chair: "Individualism in Psycho-Therapy." By E. A. Bennet.
March 23, 1938.	"The World Pictures of Children. A Method of Recording and Studying Them." By Margaret Lowenfeld.
April 27, 1938.	"Psychological Demonstration of the Drawings of a Schizophrenic Artist." By H. G. Baynes.
May 25, 1938.	Symposium on "Some Problems arising in the Adoption of Children". Papers by: S. Clement Brown, Lindsay Neustatter, E. Norman.
June 22, 1938.	"Danger Situations underlying Anxiety." By Karin Stephen.
October 26, 1938.	"Appetite in Psychological Disorder." By D. W. Winnicott.
November 23, 1938.	"Recent Tendencies in Intelligence Testing." By Wm. Stephenson.

December 14, 1938.	Symposium on "Panic and Air-Raid Precautions". Papers by: John Rickman, E. B. Strauss, John Langdon-Davies, Denis Carroll.

Education Section

January 4, 1938.	"Recent Advances in the Psychology of Young Children." By Susan Isaacs.
February 14, 1938.	"Some Psychological and Educational Disabilities of Children suffering from Minor Physical Defects or Abnormalities." By James Lumsden.
March 21, 1938.	"The Psychological Needs of Children and their Relation to Behaviour." By C. Sanders.
April 25, 1938.	"The Control of Frustration in the Development of Children from Two to Four Years." By D. May.
May 23, 1938.	"The Uses and Abuses of Mental Tests." By J. C. Hill and Wm. Stephenson.
October 17, 1938.	"Some Past and Present Difficulties in Child Guidance Theory and Practice." By R. E. Lucas.
November 28, 1938.	"The School as a Rural Community." By H. Morris.

Industrial Section

February 16, 1938.	"The Unemployment Insurance Act: Some Problem Cases." By Dorothy Scott Stokes.
March 22, 1938.	"Sidelights on Industrial Efficiency." By G. W. Goodall.
April 27, 1938.	Discussion on "What is an Industrial Psychologist?" Opened by C. S. Myers and Millais Culpin.
June 16, 1938.	"The Mutual Manifestations of Some Industrial Illnesses." By H. E. Collier.
October 20, 1938.	Address from the Chair: "The Use of Tests in Vocational Guidance." By Alec Rodger.
November 17, 1938.	"Industrial Psychology in Spain." By A. Chleusebairgue.
December 8, 1938.	"Assessment of Occupational Success." By M. B. Stott.

BRANCH MEETINGS

Session 1937–8

Scottish Branch

November 6, 1937.	(Edinburgh.) "The Perception of Velocity." By R. H. Thouless. "The New Revised Stanford-Binet Scales." By D. Kennedy-Fraser.

December 4, 1937.	(Glasgow.) "Different Modes of Seeing." By E. RUBIN. "An Experiment on Insight." By R. W. PICKFORD.
January 29, 1938.	(Edinburgh.) "The Ultimate Basis of Aesthetic Experience." By W. R. D. FAIRBAIRN. "Psychology and Art-Value." By T. E. DICKSON.
March 5, 1938.	(Glasgow Royal Mental Hospital.) "Some Observations on the Psychological Content of Psychotic Reactions." By T. F. RODGER. "Case Material illustrative of Abnormal Behaviour in Children showing Psychotic Patterns." By J. D. SUTHERLAND.
March 31–April 4, 1938.	(St Andrews.) Extended General Meeting of the Society. (Arranged with the co-operation of the Scottish Branch.) (See under GENERAL MEETINGS.)

MIDLAND BRANCH

December 15, 1937.	"Fads and Fancies in the Psychology of Early Childhood." By C. W. VALENTINE.
February 2, 1938.	"Grades of Awareness and the Unconscious." By J. LEYCESTER KING, S.J.
March 2, 1938.	"Falling Asleep and Waking." By MARY STURT.
May 14, 1938.	"The Scientific Treatment of Delinquency." By D. CARROLL and M. STUART MILLER.
October 19, 1938.	"Of What Value are Psychological Methods in our Schools To-day?" By L. BRADLEY and T. F. ROGERS.

NORTHERN BRANCH

October 9, 1937.	Meeting held in the College of Technology, Rotherham. "The Value of the Essay Paper at the Junior Scholarship Examination of a County Education Authority." By R. K. ROBERTSON. "An Investigation of the Social Status of the Homes of Children attending a School for Mental Defectives." By J. M. COHEN. "Hope and the Mental Defective." By F. J. S. ESHER.

CAMBRIDGE: PRINTED BY W. LEWIS, M.A., AT THE UNIVERSITY PRESS